

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2002, 16:46:22 : Search time 35 Seconds
(without alignments)
449.245 Million cell updates/sec

Title: US-09-598-443-2_COPY_1_118
Perfect score: 626
Sequence: 1 MPWCDBRAPFLSPEDQVL.....SIGNIFSSFTLRAGPTSH 118

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Reached: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	626	100.0	410	20	AA125426 Human SIGIRR prote
2	626	100.0	410	22	AA1981 Human NOV9 protein
3	626	100.0	504	23	AA1981 Human NOV9 protein
4	240	38.3	407	22	AA17408 Novel signal trans
5	240	38.3	407	22	AA17408 Novel signal trans
6	240	38.3	410	22	AA17408 Novel signal trans
7	234	37.4	363	21	AA158416 Human polypeptide
8	105	16.8	871	22	AA130570 Lung cancer associ
9	105	16.8	871	23	AA130570 Lung cancer associ
10	105	16.8	1351	22	AA130567 Human kinase PRIN
					Amino acid sequenc

11	105	16.8	1610	22	AA130568 A full length huma
12	105	16.8	1618	22	AA185504 Human protein kina
13	105	16.8	1665	23	AA191511 Human kinase (PRIN
14	105	16.8	1665	23	AA191511 Human myosin light
15	105	16.8	2596	22	AA130569 A full length mous
16	98	15.7	548	22	AA130571 A full length rat
17	97	15.5	548	22	AA130571 A full length rat
18	96	15.3	534	22	AA166941 Drosophila melanog
19	94.5	15.1	295	22	AA166941 Drosophila melanog
20	90.5	14.5	506	22	AA166941 Drosophila melanog
21	90.5	14.5	524	22	AA166941 Drosophila melanog
22	90.5	14.5	524	22	AA166941 Drosophila melanog
23	90.5	14.5	654	22	AA166941 Drosophila melanog
24	90.5	14.5	1694	22	AA166941 Drosophila melanog
25	90.5	14.5	1700	23	AA166941 Drosophila melanog
26	90.5	14.5	1709	22	AA166941 Drosophila melanog
27	90.5	14.5	1839	22	AA166941 Drosophila melanog
28	88	14.1	1336	22	AA166941 Drosophila melanog
29	88	14.1	1336	22	AA166941 Drosophila melanog
30	87.5	14.0	3117	21	AA166941 Drosophila melanog
31	86	13.7	898	22	AA166941 Drosophila melanog
32	84.5	13.5	739	22	AA166941 Drosophila melanog
33	84	13.4	292	22	AA166941 Drosophila melanog
34	84	13.4	315	22	AA166941 Drosophila melanog
35	84	13.4	465	19	AA166941 Drosophila melanog
36	84	13.4	937	22	AA166941 Drosophila melanog
37	84	13.4	952	22	AA166941 Drosophila melanog
38	82	13.1	394	21	AA166941 Drosophila melanog
39	82	13.1	394	21	AA166941 Drosophila melanog
40	82	13.1	628	22	AA166941 Drosophila melanog
41	82	13.1	628	22	AA166941 Drosophila melanog
42	81.5	13.0	496	22	AA166941 Drosophila melanog
43	81	12.9	816	14	AA166941 Drosophila melanog
44	80	12.8	1571	19	AA166941 Drosophila melanog
45	80	12.8	1910	19	AA166941 Drosophila melanog

ALIGNMENTS

RESULT 1	
AA125426	
AA125426 standard; Protein; 410 AA.	
AA125426;	
08-SEP-1999 (first entry)	
Human SIGIRR protein.	
SIGIRR: anti-inflammatory; anti-autoimmune disease; chromosome 11;	
interleukin-1 receptor; IL15.5; arthropodosis multiplex congenita;	
breast cancer; insulin-dependent diabetes; sickle cell anaemia;	
bladder cancer; detection; cell-signal transduction; gene therapy;	
immune regulation; cell proliferation; cell death; cell migration;	
cell interaction; inflammation; marker; mass spectrometry; human.	
OS Homo sapiens.	
XX	
PN WO932626-A1	
XX	
PD 01-JUL-1999.	
XX	
PF 23-DEC-1998: 98WO-US27368.	
XX	
PR 23-DEC-1997: 97US-0068770.	
XX	
PA (IMMUNEX) IMMUNEX CORP.	
XX	
PI Sims JE:	
XX	
DR WPI: 1999-418925/35.	
DR N-PSDB: AA186091.	

xx New interleukin-1 receptor analog SIGIRR nucleic acid and proteins

ps Claim 2; Page 69-71; 72pp; English.

cc This invention describes a novel human SIGIRR DNA, its allelic variants
cc or species homologues which have anti-inflammatory and anti-autoimmune
cc disease activity. SIGIRR is an analog of interleukin-1 receptor. The
cc products of the invention are used (a) as probes or primers for
cc identifying nucleic acid that encodes proteins with SIGIRR activity; (b)
cc to identify human chromosome 11, to map genes on this chromosome and to
cc identify disease-related genes (particularly in the region 11p15.5 where
cc genes are present associated with e.g. arthrogryposis multiplex
cc congenita, breast cancer, insulin-dependent diabetes, sickle cell
cc anaemia, bladder cancer), including detection of defective genes; (c)
cc to study cell-signal transduction and the SIGIRR system, and (d) in
cc gene therapy. Sense and antisense oligonucleotides can be used to inhibit
cc expression of the SIGIRR gene. The proteins of the invention are used:
cc (i) to study cellular processes (immune regulation, proliferation, death,
cc migration, interaction with other cells and inflammation); (ii) to
cc identify and purify proteins that associate with SIGIRR ligands and
cc receptors, and to measure their biological activity; (iii) in screening
cc for and rational design of, potential inhibitors of activity; (iv)
cc therapeutically against diseases mediated by SIGIRR polypeptide
cc counter-structures; (v) as molecular weight (m.w.) markers in
cc electrophoresis; (vi) for determining isoelectric points of unknown
cc proteins; (vii) as controls for determining the extent of protein
cc fragmentation (e.g. to aid characterization of protein structures by
cc mass spectrometry); (viii) for generation of antibodies (Ab); (ix) to
cc deliver diagnostic or therapeutic agents to cells that express SIGIRR
cc binding molecules.

xx Sequence 410 AA;

xx Query Match 100.0%; Score 626; DB 20; Length 410;

xx Best Local Similarity 100.0%; Pred. No. 1.6e-57;

xx Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPGVCDRAPDLPSEDOVLRPALGSSVALNCTAWVSGPHGSLPSVOMLKDGLPIGIG 60

Db 1 MPGVCDRAPDLPSEDOVLRPALGSSVALNCTAWVSGPHGSLPSVOMLKDGLPIGIG 60

Qy 61 HYSLHEYSWKANLSEVLYSSVLGVNVTSEYVGAFTCSIQNISFSFTLQRAQPTSH 118

Db 61 HYSLHEYSWKANLSEVLYSSVLGVNVTSEYVGAFTCSIQNISFSFTLQRAQPTSH 118

xx RESULT 2

xx AAB61139 standard; Protein: 410 AA.

xx AAB61139;

xx 30-MAR-2001 (first entry)

xx Human NOV9 protein.

xx Human: NOV9; anti-inflammatory; cytosolic; neuroprotective;

xx cerebroprotective; immunomodulator; vulnerrary; vasotropic; gene therapy;

xx hyperplasia; tumour; restenosis; psoriasis; Dupuytren's contracture;

xx diabetes; rheumatoid arthritis; cerebral oedema; Alzheimer's disease.

xx Homo sapiens.

xx WO200075321-A2

xx 14-DEC-2000.

xx 01-JUN-2000; 2000WO-US15303.

xx 03-JUN-1999; 99US-0137322.

xx 16-MAR-2000; 2000US-0189810.

xx 22-MAR-2000; 2000US-0191158.

PR 30-MAR-2000; 2000US-0193086.

PR 31-MAY-2000; 2000US-0137322.

xx (CURA-) CURAGEN CORP.

xx Shinkets RA, Fernandes E, Herrman J, Vernet C;

xx WPI: 2001-102403/11.

xx N-PSDB; AAF27857.

xx New NOV9 polypeptides and polynucleotides, useful in gene therapy, as a

xx diagnostic marker, protein therapeutic, antibody or small molecule drug

xx target for treating immune, proliferative and metabolic diseases and

xx wound healing

xx Claim 1; Page 36-38; 194pp; English.

cc The present sequence is a new isolated polypeptide (NOV9). The NOV9
cc polypeptides, NOV9 nucleic acids, and anti-NOV9 antibodies are useful for
cc treating or preventing NOV9-associated disorders. They are also useful
cc for determining the presence of or a predisposition to a disease
cc associated with altered levels of the NOV9 polypeptide or nucleic acid.
cc These NOV9-associated disorders include hyperplasia, tumours,
cc rheumatoid arthritis, cerebral lesions, diabetic neuropathies, cerebral
cc oedema, senile dementia or Alzheimer's disease. The NOV9 polynucleotides
cc are especially useful in gene therapy. Specifically, NOV9 is useful as
cc a diagnostic marker or prognostic marker, protein therapeutic and
cc antibody target or small molecule drug target to treat disorders in the
cc immune response pathway, thyroid and metabolic diseases, bone metabolic
cc disorders, diseases of the pancreas (e.g. diabetes or digestive
cc disorders), proliferative diseases, or tissue regeneration and
cc development (e.g. wound healing or treatment of burns).

xx Sequence 410 AA;

xx Query Match 100.0%; Score 626; DB 22; Length 410;

xx Best Local Similarity 100.0%; Pred. No. 1.6e-57;

xx Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPGVCDRAPDLPSEDOVLRPALGSSVALNCTAWVSGPHGSLPSVOMLKDGLPIGIG 60

Db 1 MPGVCDRAPDLPSEDOVLRPALGSSVALNCTAWVSGPHGSLPSVOMLKDGLPIGIG 60

Qy 61 HYSLHEYSWKANLSEVLYSSVLGVNVTSEYVGAFTCSIQNISFSFTLQRAQPTSH 118

Db 61 HYSLHEYSWKANLSEVLYSSVLGVNVTSEYVGAFTCSIQNISFSFTLQRAQPTSH 118

xx RESULT 3

xx AAU86151 standard; Protein: 504 AA.

xx AAU86151;

xx 15-JUL-2002 (first entry)

xx Human hRO342 polypeptide.

xx Human: PRO; benign tumour; malignant tumour; lymphoid malignancy;

xx leukaemia; neuronal disorder; stromal disorder; blastocoele disorder;

xx inflammatory disorder; immune disorder; angiogenic disorder;

xx cytosolic; neuroprotective.

xx Homo sapiens.

xx WO200153486-A1

xx 26-JUL-2001.

xx 11-FEB-2000; 2000WO-US03565.

xx 08-MAR-1999; 99WO-US05028.

PR 11-MAR-1999; 99US-123972P.
PR 11-MAY-1999; 99US-133459P.
PR 02-JUN-1999; 99WO-US12252.
PR 22-JUN-1999; 99US-140650P.
PR 20-JUN-1999; 99US-140653P.
PR 20-JUL-1999; 99US-144758P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 17-AUG-1999; 99US-149395P.
PR 31-AUG-1999; 99US-151689P.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 05-JAN-2000; 2000WO-US00219.
XX

(GETH) GENENTECH INC.

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Plettl RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX WPI: 2002-205567/26.
DR N-PSDB; ABK40277.
XX

PT Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
XX

PS Claim 61; Fig 48; 302pp; English.

XX The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
CC useful for treating benign or malignant tumours (e.g. renal, kidney,
CC bladder, breast, etc), leukemias and lymphoid malignancies, other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC macrophagal, stromal and blastocellic disorders, inflammatory, immune
CC and angiogenic disorders. The polynucleotide sequences are also
CC useful in gene therapy. AA086128-AA086162 represent the human PRO
CC polypeptides of the invention.
XX

SQ Sequence 504 AA:

Query Match 100.0%; Score 626; DB 23; Length 504;
Test Local Similarity 100.0%; Pred. No. 2e-57;
atches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MFGVCDRADPFLSPEDVLRPALSSVALNCTAMVYSGPHCSLPVOMLXKGLPIGTG 60
Db 1 MFGVCDRADPFLSPEDVLRPALSSVALNCTAMVYSGPHCSLPVOMLXKGLPIGTG 60

OY 61 HYSLEHYSWVKANLSEVLSSVLYGVNVTSEYVGAFTCSIQNISPSFTLQAGPTSH 118
Db 61 HYSLEHYSWVKANLSEVLSSVLYGVNVTSEYVGAFTCSIQNISPSFTLQAGPTSH 118

RESULT 4
ID AAU17408 standard; Protein; 407 AA.
XX
AC AAU17408;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel signal transduction pathway protein, seq ID 973.
XX
XX Neutropenic; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX Homo sapiens.
XX WO200154733-A1.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01312.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226881.
PR 22-AUG-2000; 2000US-0226886.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228927.
PR 01-SEP-2000; 2000US-0228928.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.

PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SW;
XX WPI; 2001-465460/50.
DR N-PDB; AAS27325.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders
XX
PS Claim 1; SEQ ID No 973; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders, in
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AA017059-AA017683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX
Query Match 38.3%; Score 240; DB 22; Length 407;
Best Local Similarity 100.0%; Pred. No. 7.5e-17;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 70 VKANLSEVLVSSVLYGNVTSTVEYGAFTCSIONISFSSFTIQRAGPTSH 118
Db 67 VKANLSEVLVSSVLYGNVTSTVEYGAFTCSIONISFSSFTIQRAGPTSH 115
RESULT 5
ID AAM43652
ID AAM43652 standard; Protein; 407 AA.
AC AAM43652;
XX
XX 22-OCT-2001 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 330.
DE
XX
KW Human; antiarrhythmic; antirheumatic; antiproliferative; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; vitropic;
KW fungicide; ophthalmological; cytotoxic; immunosuppressive; nootropic;
KW neuroprotective; antiarrhythmic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulvar; anticonvulsant; antibacterial;
KW antiparasitic; gardant; gene therapy; cancer; immune disorder;
KW cardiovascular disorder; neurological disease; infection; human.
OS Homo sapiens.
XX
XX
XX WO200155308-A2.
PN
XX

३

D

PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250191.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-488781/53.
DR N-PSDB: AAI63889.
XX
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders -
XX
XX
XX Claim 11: SEQ ID NO 261; 664pp + sequence listing; English.
XX
XX The invention relates to human polynucleotides (AAI63803-AAI64012) and
CC the encoded proteins (AAM44497-AAM43660) useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. The
CC genes were isolated from a range of human tissues disclosed in the
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,

CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC uterine; (b) immune disorders e.g. Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX
SQ Sequence 410 AA:
Query Match 38.3%; Score 240; DB 22; Length 410;
Best Local Similarity 100.0%; Pred. No. 7.6e-17;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 70 VKANLSEVLVSSVLGVNVTSTVEYGAFTCSIONISPSFTLRAGPTSH 118
Db 70 VKANLSEVLVSSVLGVNVTSTVEYGAFTCSIONISPSFTLRAGPTSH 118
RESULT 7
AAB58416
ID AAB58416 standard; Protein; 363 AA.
XX
AC AAB58416;
XX
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polypeptide sequence SEQ ID 754.
XX
XX Homo sapiens.
OS
XX WO20005180-A2.
PN
XX 21-SEP-2000.
PD
XX 08-MAR-2000; 2000WO-US05918.
PF
XX 12-MAR-1999; 99US-0124270.
PR
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX Ruben SM;
XX
XX WPI: 2000-587514/55.
DR N-PSDB: AAI6292.
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX
XX
XX Claim 11: Page 1271-1272; 1425pp; English.
XX
XX Polynucleotide sequences AAI7982 - AAI8424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective, cytoprotective, cardioprotective;
CC immunomodulatory; muscular active general; vulnerrary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other

CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of tumors and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC AAB38349 are used in the course of the invention for the
 CC identification and characterization of the polynucleotide and protein
 CC sequences.

XX Sequence 363 AA;

Query Match 37.4%; Score 234; DB 21; Length 363;
 Best Local Similarity 98.0%; Pred. No. 2.8e-16;
 Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 70 VKANLSEVLYVSLGVNVTSTEVYGAFTCSIONISFSSFTLRAGPTSH 118
 DB 23 VKANLSEVLYVSLGVNVTSTEVYGAFTCSIONISFSSFTLRAGPTSH 71

RESULT 8

00570 AAB30570 standard; Protein: 871 AA.

XX AAB30570;

XX 19-MAR-2001 (first entry)

DE A splice variant of a signal transduction polypeptide.

XX Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
 KW congestive heart failure; dilated congestive cardiomyopathy;
 KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
 KW mitral valve disease; aortic valve disease; tricuspid valve disease;
 KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;
 KW atherosclerosis; cardiac tumour; microbial infection; splice variant.

OS Homo sapiens.

XX WO200063381-A1.

XX 26-OCT-2000.

XX 11-APR-2000; 2000WO-US09488.

XX 16-APR-1999; 99US-0129553.

XX (SCIO-) SCIOS INC.

XX Zeng W, Stanton L, Kong H;

XX WPI; 2001-007013/01.

XX Claim 1; Page 74-76; 81pp; English.

CC The present sequence represents a splice variant of human in signal
 CC transduction polypeptide. The polypeptide is designated H19G5. The
 CC protein is capable of regulating signal transduction and exhibits kinase
 CC activity. The H19G5 transcript is expressed in the heart. H19G5
 CC polypeptides and polynucleotides are useful for preventing or treating a
 CC cardiac disease, such as congestive heart failure, dilated congestive
 CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
 CC mitral valve disease, aortic valve disease or tricuspid valve disease,
 CC angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
 CC arterial or rheovascular hypertension, arteriosclerosis, atherosclerosis,
 CC and cardiac tumours in humans. The polypeptide is also useful for
 CC detecting the expression of a protein capable of regulating signal
 CC transduction or the expression of a protein capable of acting as a donor
 CC or acceptor molecule of a phosphate group. The monoclonal antibodies can

CC be used as probes for detecting discrete antigens expressed by tissue or
 CC cell samples, and therefore used in humans for localization and
 CC monitoring of microbial infection.

XX Sequence 871 AA;

Query Match 16.8%; Score 105; DB 22; Length 871;
 Best Local Similarity 31.4%; Pred. No. 0.034;
 Matches 33; Conservative 11; Mismatches 29; Indels 32; Gaps 5;
 6 DRAPFLSPSPDQVLPALGSSVALNCTAWVSGPCLPSVO-WKQDLPLIGIHS 63
 DB 363 DRAPFLRELSDERY-VLGOSVTLACV-----SAQPAQATWSKDGAPL----- 406
 64 LHEYSWKRANLSEVLYS-----VLGVNVTSTEVYGAFTCSION 102
 DB 407 -----ESSSRVLTLSATLKNPQLRLIIVVAEDLVITCVSN 443

RESULT 9

AAE16274 AAE16274 standard; Protein: 871 AA.

XX AAE16274;

XX 26-MAR-2002 (first entry)

DE Human kinase PKIN-20 protein.

XX Human; kinase; PKIN-20; cancer; leukemia; adenocarcinoma; osteoporosis;
 KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
 KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anemia;
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;
 KW Good pasture's syndrome; Graves' disease; pancreatitis; porphyria;
 KW Rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
 KW cardiovascular disease; hypertension; vascular disease; myocarditis; obesity;
 KW congestive heart failure; ischemic heart disease; lung tumour; gout;
 KW fatty liver; Niemann-Pick's disease; gene therapy.

OS Homo sapiens.

XX Key Location/Qualifiers

XX Domain /note="Eukaryotic protein kinase domain"

XX Domain /label="Protein_kinase_domain"

XX WO200196547-A2.

XX 20-DEC-2001.

XX 14-JUN-2001; 2001WO-US19444.

XX 15-JUN-2000; 2000US-212073P.

XX 23-JUN-2000; 2000US-213467P.

XX 30-JUN-2000; 2000US-213651P.

XX 07-JUL-2000; 2000US-213651P.

XX 13-JUL-2000; 2000US-213651P.

XX 25-AUG-2000; 2000US-228056P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;

PI Gandhi AR, Trifunovic CA, Walla NK, Yao MG, Lu DM, Greenwald SR;

PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YF;

PI Baughman MR, He A, Thornton M, Hafila A, Patterson C, Gururajan R;

PI Lo TP, Khan F, Reichenbach SA, Azimzai Y, Policky JL, Ding L;

PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;

XX WPI; 2002-090207/12.

XX N-PSDB; AAD26467.

XX New polypeptides, useful for diagnosing, treating or preventing
PT disorders of growth and development, cardiovascular and lipid, and
PT diseases such as cancer, comprise human kinase polypeptides -

Claim 1: Page 164-165; 197pp; English.

XX The invention relates to human kinase PKIN proteins and their
CC corresponding cDNAs. A composition containing PKIN agonist is useful for
CC treating a disease or condition associated with decreased expression of
CC PKIN and a composition comprising PKIN antagonist is useful for treating
CC a disease or condition associated with overexpression of PKIN. The
CC disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,
CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
CC atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
CC growth and development disorders (arteriosclerosis, cataracts); Cardio
CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); Cardio
CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
CC hypocholesterolaemia, obesity). PKIN DNA is useful for assessing
CC toxicity of a test compound and in gene therapy. The present sequence
CC is human PKIN-20 protein.

XX Sequence 871 AA:

Query Match 16.8%; Score 105; DB 23; Length 871;
Best Local Similarity 31.4%; Pred. No. 0.034;
Matches 33; Conservative 11; Mismatches 29; Indels 32; Gaps 5;

OY 6 DRAPDLFSEDDOVLRLPALGSSVALNCTAMWVSGPHCSLPVQ--WLKGLPLGIGHYS 63

DB 363 DRAPFLRLSLDETV--VLGQSVTLACQV-----SAQPAQATWSKQDAPL----- 406

OY 64 LHEYSWKANLSEVLVS-----VLGVNVTSTEVYGAFTCSION 102

DB 407 -----ESSSRVLISATLKNQQLTLVVAEDGYVTCVSN 443

RESULT 10

ID AAB30567 standard; Protein: 1351 AA.

XX AAB30567;

DT 19-MAR-2001 (first entry)

DE Amino acid sequence of a human signal transduction polypeptide.

XX Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
KM congestive heart failure; dilated congestive cardiomyopathy;
KM hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
KM mitral valve disease; aortic valve disease; tricuspid valve disease;
KM myocardial infarction; cardiac arrhythmia; arteriosclerosis;
KM atherosclerosis; cardiac tumour; microbial infection.

XX Homo sapiens.

PN WO200063381-A1.

PD 26-OCT-2000.

PF 11-APR-2000; 2000WO-US09488.

PR 16-APR-1999; 99US-0129553.

XX

PA (SCIO-) SCIOS INC.

PI Zeng W, Stanton L, Kong H;

DR WPI: 2001-007013/01.

DR N-PSDB: AAC62285.

PT Novel h19G5 polypeptides capable of regulating signal transduction and
PT exhibiting kinase activity useful for identifying antibodies to treat
PT cardiac diseases, and additional mediators of signal transduction -

PS Claim 1: Page 55-57; 81pp; English.

XX The present sequence represents a human protein with putative function
CC in signal transduction. The polypeptide is designated H19G5. The protein
CC is capable of regulating signal transduction and exhibits kinase
CC activity. The H19G5 transcript is expressed in the heart. H19G5
CC polypeptides and polynucleotides are useful for preventing or treating a
CC cardiac disease, such as congestive heart failure, dilated congestive
CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
CC mitral valve disease, aortic valve disease or tricuspid valve disease,
CC angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
CC arterial or renovascular hypertension, arteriosclerosis, atherosclerosis
CC and cardiac tumours in humans. The polypeptide is also useful for
CC detecting the expression of a protein capable of regulating signal
CC transduction or the expression of a protein capable of acting as a donor
CC or acceptor molecule of a phosphate group. The monoclonal antibodies can
CC be used as probes for detecting discrete antigens expressed by tissue or
CC cell samples, and therefore used in humans for localization and
CC monitoring of microbial infection.

XX Sequence 1351 AA:

Query Match 16.8%; Score 105; DB 22; Length 1351;
Best Local Similarity 31.4%; Pred. No. 0.06;
Matches 33; Conservative 11; Mismatches 29; Indels 32; Gaps 5;

OY 6 DRAPDLFSEDDOVLRLPALGSSVALNCTAMWVSGPHCSLPVQ--WLKGLPLGIGHYS 63

DB 843 DRAPFLRLSLDETV--VLGQSVTLACQV-----SAQPAQATWSKQDAPL----- 886

OY 64 LHEYSWKANLSEVLVS-----VLGVNVTSTEVYGAFTCSION 102

DB 887 -----ESSSRVLISATLKNQQLTLVVAEDGYVTCVSN 923

RESULT 11

ID AAB30568 standard; Protein: 1610 AA.

XX AAB30568;

DT 19-MAR-2001 (first entry)

DE A full length human signal transduction polypeptide.

XX Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
KM congestive heart failure; dilated congestive cardiomyopathy;
KM hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
KM mitral valve disease; aortic valve disease; tricuspid valve disease;
KM myocardial infarction; cardiac arrhythmia; arteriosclerosis;
KM atherosclerosis; cardiac tumour; microbial infection.

XX Homo sapiens.

PN WO200063381-A1.

PD 26-OCT-2000.

PF 11-APR-2000; 2000WO-US09488.

PR 16-APR-1999; 99US-0129553.

XX

PA (SCIO-) SCIOS INC.
 XX
 PI Zeng W, Stanton L, Kong H;
 XX
 DR WPI: 2001-007013/01.
 DR N-PSDB; AAC62286.
 XX
 PT Novel h19c5 polypeptides capable of regulating signal transduction and
 PT exhibiting kinase activity useful for identifying antibodies to treat
 PT cardiac diseases, and additional mediators of signal transduction
 XX
 PS Claim 1: Page 61-65; 81pp; English.
 XX
 CC The present sequence represents a human protein with putative function
 CC in signal transduction. The polypeptide is designated h19c5. The protein
 CC is capable of regulating signal transduction and exhibits kinase
 CC activity. The h19c5 transgene is expressed in the heart. h19c5
 CC polypeptides and polynucleotides are useful for preventing or treating
 CC cardiac disease, such as congestive heart failure, dilated congestive
 CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
 CC mitral valve disease, aortic valve disease or tricuspid valve disease,
 CC arterial or rheovascular hypertension, cardiac arrhythmia, pulmonary,
 CC and cardiac tumors in humans. The polypeptide is also useful for
 CC detecting the expression of a protein capable of regulating signal
 CC transduction or the expression of a protein capable of acting as a donor
 CC or acceptor molecule of a phosphate group. The monoclonal antibodies can
 CC be used as probes for detecting discrete antigens expressed by tissue or
 CC cell samples, and therefore used in humans for localization and
 CC monitoring of microbial infection.
 XX
 SQ Sequence 1610 AA;
 XX
 Query Match 16.8%; Score 105; DB 22; Length 1610;
 Best Local Similarity 31.4%; Pred. No. 0.075; 29; Indels 32; Gaps 5;
 Matches 33; Conservative 11; Mismatches 29; Indels 32; Gaps 5;
 XX
 OY 6 DRAPDLSPEDVLRPAIGSSVALNCTAMVYSGPHCSLPVQ--WLKDGPLIGIGHYS 63
 DB 1102 DRAPFLRELSDERY--VLGOSVTLACV-----SAQPAQAATWSKDGAPL----- 1145
 OY 64 LHEYSWKANLSEVLYSS-----VLGVNVTSTEVGAPFCISQION 102
 DB 1146 -----ESSSRVLISATLKNFOLLTLIVVAEDLVYTCVSNN 1182
 XX
 RESULT 12
 AAB85504
 ID AAB85504 standard; protein; 1618 AA.
 XX
 AAB85504;
 XX
 DT 25-SEP-2001 (first entry)
 DE
 XX
 DE Human protein kinase SGK145.
 XX
 KW Protein kinase; enzyme; cytosolic; neurotrophic; neuroprotective; human;
 KW antiparkinsonian; virucide; antibacterial; antifungal; antimigraine;
 KW analgesic; hypotensive; hypertensive; immunosuppressive; anti-allergic;
 KW anxiolytic; antithrombotic; antiarrhythmic; ophthalmologic; anorectic;
 KW osteopathic; thrombolytic; antiarteriosclerotic; antidiabetic; gene therapy.
 KW
 XX
 OS Homo sapiens.
 XX
 OS WO200155356-A2.
 XX
 PD 02-AUG-2001.
 XX
 XX 25-JAN-2001; 2001WO-US02337.
 XX
 PR 25-JAN-2000; 2000US-0178078.
 PR 31-JAN-2000; 2000US-0179364.

PR 17-FEB-2000; 2000US-0183173.
 PR 17-MAR-2000; 2000US-0190162.
 PR 29-MAR-2000; 2000US-0193404.
 PR 13-NOV-2000; 2000US-0247013.
 XX
 PA (SUGEN-) SUGEN INC.
 XX
 PI Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;
 XX
 DR WPI: 2001-476202/51.
 DR N-PSDB; AAH46904.
 XX
 PT Kinase polypeptides useful for treating cancers, Alzheimer's disease,
 PT viral infections, diabetes, obesity, organ transplant rejection and
 PT rheumatoid arthritis.
 XX
 PS Claim 7: Page 215; 218pp; English.
 XX
 CC The invention provides human protein kinases and protein kinase-like
 CC enzymes and polynucleotides encoding the polypeptides. The kinase
 CC polypeptides and their modulators are useful for treating a disease or
 CC disorder such as cancer, immune-related diseases, cardiovascular disease,
 CC brain or neuronal-associated disease and metabolic disorders, including
 CC cancers of tissues, cancers of hematopoietic origin, diseases of the
 CC central nervous system, diseases of the peripheral nervous system,
 CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 CC lateral sclerosis, viral infections, infections caused by prions,
 CC mood, and lung, ocular diseases, migraines, pain, sexual dysfunction,
 CC hypothyroidism, attention disorders, neurological disorders, dyskinesias,
 CC hyperthyroidism, psychotic disorders, neurodegenerative disorders,
 CC osteoarthritis, arthritis, autoimmunity, atherosclerosis, psoriasis,
 CC for treating diabetes, and organ transplant rejection. They are also useful
 CC for treating chronic inflammatory diseases such as reperfusion
 CC injury, coronary thrombosis, clotting disorders and atherosclerosis,
 CC ocular diseases such as glaucoma, retinopathy and macular degeneration,
 CC psychiatric and neurological disorders such as anxiety, schizophrenia,
 CC dementia, manic depression, etc. The polynucleotides are useful in gene
 CC therapy techniques to treat the above mentioned disorders. Sequences
 CC AAB85491-85522 represent the human protein kinases of the invention.
 XX
 SQ Sequence 1618 AA;
 XX
 Query Match 16.8%; Score 105; DB 22; Length 1618;
 Best Local Similarity 31.4%; Pred. No. 0.075; 29; Indels 32; Gaps 5;
 Matches 33; Conservative 11; Mismatches 29; Indels 32; Gaps 5;
 XX
 OY 6 DRAPDLSPEDVLRPAIGSSVALNCTAMVYSGPHCSLPVQ--WLKDGPLIGIGHYS 63
 DB 1110 DRAPFLRELSDERY--VLGOSVTLACV-----SAQPAQAATWSKDGAPL----- 1153
 OY 64 LHEYSWKANLSEVLYSS-----VLGVNVTSTEVGAPFCISQION 102
 DB 1154 -----ESSSRVLISATLKNFOLLTLIVVAEDLVYTCVSNN 1190
 XX
 RESULT 13
 AAE24151
 ID AAE24151 standard; protein; 1665 AA.
 XX
 AAE24151;
 XX
 DT 23-SEP-2002 (first entry)
 DE
 XX
 DE Human kinase (PKIN)-22 protein.
 XX
 KW Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
 KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
 KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
 KW development; hepatitis; cardiovascular; hypertension; drug screening;
 KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
 KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;

KW hypercholesterolaemia; obesity; gene therapy; cytostatic; anti-HIV;
 KW neuroprotective; hepatotropic; hypotensive; cardiast; nephrotropic;
 KW hyperlipidaemia; enzyme.
 OS Homo sapiens.
 Key Location/Qualifiers
 FH Domain 68..128 /note="Immunoglobulin domain"
 FT Domain 165..418 /note="Eukaryotic protein kinase domain"
 FT Domain 167..401 /note="Protein kinase domain"
 FT Domain 1174..1235 /note="Immunoglobulin domain"
 FT Domain 1369..1621 /note="Eukaryotic protein kinase domain"
 FT Domain 1372..1606 /note="Protein kinase domain"
 Domain
 PN WO200233099-A2.
 PD 25-APR-2002.
 XX 20-OCT-2001; 2001WO-US47728.
 XX 20-OCT-2000; 2000US-242410P.
 PR 27-OCT-2000; 2000US-244068P.
 PR 03-NOV-2000; 2000US-245708P.
 PR 09-NOV-2000; 2000US-247672P.
 PR 16-NOV-2000; 2000US-249565P.
 PR 22-NOV-2000; 2000US-252730P.
 PR 01-DEC-2000; 2000US-250607P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Gururajan R, Baughn MR, Walla NK, Elliott VS, Xu Y, Arvizu C;
 XX Yao MG, Rankumar J, Ding L, Tang YT, Hatalla AZA, Nguyen DB;
 PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM;
 PI Lal PG, Recipon SA, Lu DM, Borowsky ML, Thornton M, Swarnaker A;
 XX Thangavelu K, Khan FA, Ison CH;
 XX MPI: 2002-454603/48.
 DR N-PSDB: AAD38865.
 XX New human kinase polypeptide, for diagnosing, preventing and treating
 cancer, immune system disorders, growth and development disorders,
 cardiovascular disorders and lipid disorders
 PS Claim 1; Page 182-186; 210pp; English.
 XX The invention relates human kinases (PKIN) and their corresponding
 CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
 CC treating and preventing cancer, an immune system disorder (e.g.,
 CC acquired immune deficiency syndrome (AIDS), Addison's disease, allergy,
 CC asthma, atherosclerosis, multiple sclerosis, psoriasis), disorders
 CC affecting growth and development (e.g., arteriosclerosis, cirrhosis,
 CC hepatitis), cardiovascular disorder (e.g., hypertension, myocardial
 CC infarction, Goodpasture's syndrome), and a lipid disorder (e.g., fatty
 CC liver, Gaucher's disease, Niemann-Pick's disease, hypercholesterolaemia,
 CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
 CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
 CC condition or a disease associated with the expression of PKIN in a
 CC biological sample. A composition comprising PKIN or an agonist or
 CC antagonist of PKIN is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional PKIN.
 CC PKIN is useful in a number of drug screening techniques and to analyse
 CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
 CC knockin humanised animals or transgenic animals to model human diseases,
 CC and in somatic or germline gene therapy. The present sequence is human
 CC PKIN protein.
 XX Sequence 1665 AA;
 SQ

Query Match 16.8%; Score 105; DB 23; Length 1665;
 Best Local Similarity 31.4%; Pred. No. 0.078; Indels 32; Gaps 5;
 Matches 33; Conservative 11; Mismatches 29;
 QY 6 DRAPDFLSPEDQVLRPALGSSVALNCTAWVVGSPHCSLPSVQ--WLKDGFLPGIGGHS 63
 DB 1157 DRAPFTLRELSDFTV--VLGGSVTLACQV-----SAQPAQATWSKQCAPL----- 1200
 QY 64 LHEYSWKANLSEVLVS-----VLGVNVTSEVYGAFTCSTQN 102
 DB 1201 -----ESSSRVLISATLKNFQLTLTVVVAEDIGVYTCVSVN 1237
 RESULT 14
 AAO15372
 ID AAO15372 standard; Protein; 1665 AA.
 AC AAO15372;
 XX 19-SEP-2002 (first entry)
 DE Human myosin light chain kinase subfamily-related kinase protein.
 XX Human: gene therapy; chromosome 1; kinase protein;
 KW myosin light chain kinase subfamily; kinase protein-mediated disease;
 KW transgenic animal.
 XX Homo sapiens.
 OS WO200240683-A2.
 PN 23-MAY-2002.
 PD 22-OCT-2001; 2001WO-US32616.
 PF 14-NOV-2000; 2000US-0711134.
 PR 17-MAY-2001; 2001US-0856664.
 XX (PEKE) PE CORP NY.
 PA Wei M, Ketchum K, Di Francesco V, Beasley EM;
 XX MPI: 2002-500223/53.
 DR N-PSDB: AAL43908, AAL43909.
 XX New kinase proteins related to myosin light chain kinase subfamily and
 encoding polynucleotide, useful for diagnosing, treating disease or
 PT condition mediated by the kinase protein and for identifying modulators
 PT
 PS Claim 1; Fig 2; 96pp; English.
 XX The invention comprises the amino acid and coding sequences (located on
 CC chromosome 1) of a human kinase protein that is related to the myosin
 CC light chain kinase subfamily. The human kinase DNA and protein sequences
 CC of the invention are useful for identifying agents that modulate the
 CC activity of the human kinase protein. Kinase-modulating agents are useful
 CC for treating a disease or condition mediated by a human kinase protein.
 CC The human kinase DNA sequences can be used to produce transgenic animals
 CC which are useful for studying the function of kinase proteins and
 CC identifying/evaluating modulators of kinase protein activity. The present
 CC amino acid sequence represents the human kinase protein of the invention.
 XX Sequence 1665 AA;
 SQ
 Query Match 16.8%; Score 105; DB 23; Length 1665;
 Best Local Similarity 31.4%; Pred. No. 0.078; Indels 32; Gaps 5;
 Matches 33; Conservative 11; Mismatches 29;
 QY 6 DRAPDFLSPEDQVLRPALGSSVALNCTAWVVGSPHCSLPSVQ--WLKDGFLPGIGGHS 63
 DB 1157 DRAPFTLRELSDFTV--VLGGSVTLACQV-----SAQPAQATWSKQCAPL----- 1200

A:Accession: T27935
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7160 <WIL>
A:Cross-references: EMBL:Z73897; PIDN:CAA98065.1; GSPDB:GN00022; CESP:ZK617.1b
A:Experimental source: clone ZK617
R:Harris, B.
submitted to the EMBL Data Library, May 1996
A:Accession: T28031
A:Reference number: Z20458
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7160 <WIL>
A:Cross-references: EMBL:Z73899; PIDN:CAA98082.1; GSPDB:GN00022; CESP:ZK617.1b
A:Experimental source: clone ZK829
C:Genetics:
A:Gene: CESP:ZK617.1b
A:Map position: 4
A:Introns: 10/3; 61/3; 135/2; 168/3; 256/2; 379/3; 405/2; 463/1; 508/3; 542/3; 574/3; 597/1; 3141/3; 3269/1; 6473/3; 7012/3; 7097/1; 7129/3
C:Superfamily: twitchin; fibronectin type III repeat homology; immunoglobulin homology;

Query Match 14.4%; Score 90; DB 2; Length 7160;
Best Local Similarity 24.0%; Pred. No. 8.4;
Matches 23; Conservative 21; Mismatches 40; Indels 12; Gaps 3;

OY 7 RAPDLPSPEDQVLRPALGSSVALNCTAMVWSGPHCSLPVQWMLDGLPLGIGHYSLHE 66
DB 101 KQPSFV--GKPRIPKDGALVMECKV-----KSASFPVAKMDGVLSKGLYHA-- 151

OY 67 YSWKANISEVLSSVLTGVNVTSTEVYGAFTCSIONISSPSS 102
DB 152 ---IFSDLGDTYLCQLEIRGSSSDAGCYRCNIRN 184

RESULT 6
T20992
hypothetical protein F15G9.4a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
A:Accession: T20992; T24733
R:Sulston, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19355
A:Accession: T20992
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5175 <WIL>
A:Cross-references: EMBL:Z47068; PIDN:CAA87335.1; GSPDB:GN00028; CESP:F15G9.4a
A:Experimental source: clone F15G9
R:Kershaw, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19929
A:Accession: T24733
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5175 <WIL>
A:Cross-references: EMBL:Z47070; PIDN:CAA87344.1; GSPDB:GN00028; CESP:F15G9.4a
A:Experimental source: clone T09B9
C:Genetics:
A:Gene: CESP:F15G9.4a
A:Map position: X
A:Introns: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3; 2512/2; 2593/3; 2699/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4838/1; 4879/1; 4941/1; 5011/1; 5077/1

Query Match 14.0%; Score 87.5; DB 2; Length 5175;
Best Local Similarity 28.1%; Pred. No. 10;
Matches 32; Conservative 15; Mismatches 38; Indels 29; Gaps 6;

OY 1 MPG-----VCDRAPDLPSPEDQVLRPALGSSVALNCTAMVWSGPHCSLPVQWMLDG 53
DB 4555 MGPFRSSRTVLHHAPOFIVKPKNTT--AATGAIIVELRCSA--AGPPH---PTIWMADG 4607

OY 54 IPLGIGHYSLHEYSWKANISEVLSSVLTGVNVTSTEVYGAFTCSIONISSPSS 107
DB 4608 -----KLIEDSKFEIAYSH-LKVTLNSTSDSGERTCMAONSVSS 4646

RESULT 7
T43290
hemiscitin precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
A:Accession: T43290; T20993; T24734
R:Voegel, B.E.; Hedgecock, E.M.
submitted to the EMBL Data Library, June 1998
A:Description: Hemiscitin is required for hemidesmosome mediated cell adhesion and ge
A:Reference number: Z22396
A:Accession: T43290
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5198 <VOG>
A:Cross-references: EMBL:AF074901; PIDN:AAC26792.1
R:Sulston, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19355
A:Accession: T20993
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5198 <WIL>
A:Cross-references: EMBL:Z47068; PIDN:CAA87336.1; GSPDB:GN00028; CESP:F15G9.4b
A:Experimental source: clone F15G9
R:Kershaw, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19929
A:Accession: T24734
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5198 <WIL>
A:Cross-references: EMBL:Z47070; PIDN:CAA87345.1; GSPDB:GN00028; CESP:F15G9.4b
A:Experimental source: clone T09B9
C:Genetics:
A:Gene: him-4; F15G9.4b
A:Map position: X
A:Introns: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3; 2512/2; 2593/3; 2699/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4838/1; 4902/1; 4964/1; 5034/1; 5177/1

Query Match 14.0%; Score 87.5; DB 2; Length 5198;
Best Local Similarity 28.1%; Pred. No. 10;
Matches 32; Conservative 15; Mismatches 38; Indels 29; Gaps 6;

OY 1 MPG-----VCDRAPDLPSPEDQVLRPALGSSVALNCTAMVWSGPHCSLPVQWMLDG 53
DB 4555 MGPFRSSRTVLHHAPOFIVKPKNTT--AATGAIIVELRCSA--AGPPH---PTIWMADG 4607

OY 54 IPLGIGHYSLHEYSWKANISEVLSSVLTGVNVTSTEVYGAFTCSIONISSPSS 107
DB 4608 -----KLIEDSKFEIAYSH-LKVTLNSTSDSGERTCMAONSVSS 4646

RESULT 8
AE3507
amidotransferase hsh (EC 2.4.2.-) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
A:Accession: AE3507
R:DelVecchio, V.G.; Kaprali, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanov, M.; Mazur, M.; Goltsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Let, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: ADJ3252; PMID:11756688
A:Accession: AE3507
A:Status: preliminary
A:Molecule type: DNA

A:Molecule type: mRNA
A:Residues: 465-518 <LA1>
A:Experimental source: sciatic nerve
C:Genetics:
A:Gene: FGFR4; tyro-9
C:Function:
A:Description: receptor mediating effects of fibroblast growth factor
A:Note: expressed in normal lung; expressed in some carcinomas
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F:11-72/Domain: immunoglobulin homology <IM1>
F:110-181/Domain: immunoglobulin homology <IM2>
F:218-238/Domain: transmembrane #status predicted <TM>
F:239-650/Domain: intracellular #status predicted <INT>
F:313-598/Domain: protein kinase homology <KIN>
F:321-329/Region: protein kinase ATP-binding motif
F:104,116,157,168/Binding site: carbonylstrate (Asn) (covalent) #status predicted
F:351,368,460/Active site: Lys, Glu, Asp #status predicted
F:55,478/Binding site: magnesium (Asn, Asp) #status predicted
F:81/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 13.0%; Score 81.5; DB 1; Length 650;
Best Local Similarity 31.9%; Pred. No. 3.6;
Matches 29; Conservative 11; Mismatches 36; Indels 15; Gaps 4;

Oy 24 LGSSVALNCTAMVYSGPHCSLPYQWLK----DGLPLIGGHYSLHEYSWKANLSEVLV 79
Db 109 VGSNVELCKKYSDQPH-----IQWLKHVYNGSSLGADGPPYQVLKTTDINSEV-- 161
Oy 80 SSVLGVNVTSTEVYGAFTCSION--ISFSS 107
Db 162 -EVLIRNVSAEDAGEYTCLAGNSIGLSYQS 191

RESULT 13
Ig Lambda-like chain, V-C region - nurse shark
A:Accession: A49633
C:Species: Ginglymostoma cirratum (nurse shark)
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
C:Accession: A49633
R:Greenberg, A.S.; Steiner, L.; Kasahara, M.; Flajnik, M.F.
Proc. Natl. Acad. Sci. U.S.A. 90, 10603-10607, 1993
A:Title: Isolation of a shark immunoglobulin light chain cDNA clone encoding a protein r
A:Reference number: A49633; MUID:94068449; PMID:8248152
A:Accession: A49633
A:Status: preliminary
Molecule type: mRNA
A:Residues: 1-238 <GRE>
Experimental source: spleen
A:Note: sequence extracted from NCBI backbone (NCBI:141012, NCBI:141013)
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:36-110/Domain: immunoglobulin homology <IM>

Query Match 12.9%; Score 81; DB 2; Length 238;
Best Local Similarity 29.3%; Pred. No. 1.2;
Matches 29; Conservative 18; Mismatches 36; Indels 16; Gaps 3;

Oy 12 LSPSEDOVLRPALGSSVALNCTAMVYSGPHCSLP--SVQWLKDGPLIGGHYSLHEYS 68
Db 141 LPPSPDOVOTKG-----TATLVCLANHFYDELQVQWKKGDAVSDGVOTS-----N 187
Oy 69 WYKANLSEVLVSSVLGVNVTSTEVYGAFTCSIONISFSS 107
Db 188 YLRASDSTYSVSLTLTSGSDMESNARPSCAITHTLSS 226

RESULT 14
S18209
fibroblast growth factor receptor 4 precursor (clone 61) - mouse
N:Alternate names: tyrosine kinase Mpk-11
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Mus musculus (house mouse)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 16-Jul-1999

C:Accession: S18209; S26751; S30497
R:Stark, K.L.; McMahon, J.A.; McMahon, A.P.
Development 113, 641-651, 1991
A:Title: FGFR-4, a new member of the fibroblast growth factor receptor family, expres
A:Reference number: S18209; MUID:92146274; PMID:1723680
A:Accession: S18209
A:Molecule type: mRNA
A:Residues: 1-799 <STRA1>
A:Cross-references: EMBL:X59927
R:Stark, K.L.
submitted to the EMBL Data Library, May 1991
A:Reference number: S26751
A:Accession: S26751
A:Molecule type: mRNA
A:Residues: 1-485, 'QVYRAAEFC', 486-799 <STRA2>
A:Cross-references: EMBL:X59927; NID:950968; PIDN:CAA42551.1; PID:950969
R:Giliardi-Hebenstreit, P.; Nieto, M.A.; Frain, M.; Mattei, M.G.; Chestler, A.; Wilkin
Oncogene 7, 2499-2506, 1992
A:Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed i
A:Reference number: S30496; MUID:93096484; PMID:1281307
A:Accession: S30497
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 611-667 <GII>
A:Cross-references: EMBL:X57236; NID:953187; PIDN:CAA40512.1; PID:953188
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: ATP; duplication; glycoprotein; growth factor receptor; phosphotransferas
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-799/Product: fibroblast growth factor receptor 4 #status predicted <ANT>
F:19-366/Domain: extracellular #status predicted <EXT>
F:162-223/Domain: immunoglobulin homology <IM>
F:367-387/Domain: transmembrane #status predicted <TM>
F:388-799/Domain: intracellular #status predicted <INT>
F:462-747/Domain: protein kinase homology <KIN>
F:470-478/Region: protein kinase ATP-binding motif
F:54-98, 169-221, 268-330/Disulfide bonds: #status predicted
F:500, 517, 609/Active site: Lys, Glu, Asp #status predicted

Query Match 12.9%; Score 80.5; DB 2; Length 799;
Best Local Similarity 30.8%; Pred. No. 5.7;
Matches 28; Conservative 11; Mismatches 37; Indels 15; Gaps 4;

Oy 24 LGSSVALNCTAMVYSGPHCSLPYQWLK----DGLPLIGGHYSLHEYSWKANLSEVLV 79
Db 260 VGSNVELCKKYSDQPH-----IQWLKHVYNGSSFGADGPPYQVLKTTDINSEV-- 312
Oy 80 SSVLGVNVTSTEVYGAFTCSION--ISFSS 107
Db 313 -QVLIRNVSAEDAGEYTCLAGNSIGLSYQS 342

RESULT 15
S20901
titin - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
C:Accession: S20901; I46520
R:Labelt, S.; Gautel, M.; Lakey, A.; Trinick, J.
EMBO J. 11, 1711-1716, 1992
A:Title: Towards a molecular understanding of titin.
A:Reference number: S20897; MUID:92258380; PMID:1582406
A:Accession: S20901
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-6805 <LAB>
A:Cross-references: EMBL:X64696
R:Labelt, S.; Barlow, D.P.; Gautel, M.; Gibson, T.; Holt, J.; Hsieh, C.L.; Francke, U
Nature 345, 273-276, 1990
A:Title: A regular pattern of two types of 100-residue motif in the sequence of titin
A:Reference number: I46520; MUID:90238553; PMID:2129545
A:Accession: I46520
A:Status: translated from GR/EMBL/DBJ

A: Molecule type: mRNA
 A: Residues: 4235-5250 <LA2>
 A: Cross-References: EMBL:X17329; NID:q1756; PID:CAA35207.1; PID:g930251
 C: Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; ptc
 C: Keywords: muscle

Query Match 12.9%; Score 80.5; DB 2; Length 6805;
 Best Local Similarity 24.4%; Pred. No. 70;
 Matches 21; Conservative 16; Mismatches 30; Indels 19; Gaps 3;

QY 36 VYSGRHCSLPYQWLKDLPLGIGSHYSLHEYSWKANISEVLVSSVVGAVNTSTEVYGA 95
 DB 6197 VIGRR-----FEIFWVKDGEPL-----KOTTRVNVETATSTILHIKSSKDDFGK 6243
 QY 96 FTCSIONT-----SFSSFTIORACP 115
 DB 6244 YTTATNSAGTATENLSYIVLEKFGP 6269

Search completed: November 14, 2002, 17:31:48
 Time: 24 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2002, 16:48:52 : Search time 13 Seconds
(without alignments)
376.477 Million cell updates/sec

Title: US-09-598-443-2_COPY_1_118
Perfect score: 626
Sequence: 1 MFCVCDRAPFLSPSEDOVL.....SIGNISFSTLQKAGPTSH 118

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Reached: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwisProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90.5	14.5	1709	1 SN_HUMAN	Q99222 homo sapien
2	83.5	13.3	216	1 HRS5_BRUCE	O8Y235 bruceella me
3	83.5	13.3	1694	1 SN_MOUSE	O62230 mus musculu
4	82	13.1	628	1 LU_HUMAN	P50895 homo sapien
5	80.5	12.9	808	1 FGR4_MOUSE	Q03142 mus musculu
6	80	12.8	2012	1 DSCA_HUMAN	O60469 homo sapien
7	79.5	12.7	1242	1 NPHN_MOUSE	O94257 mus musculu
8	78.5	12.5	1234	1 NPHN_RAT	O94044 rattus norv
9	76.5	12.2	802	1 FGR4_HUMAN	P22455 homo sapien
10	76.5	12.2	1142	1 MYRF_HUMAN	Q14324 homo sapien
11	76.5	12.2	1241	1 NPHN_HUMAN	O60500 homo sapien
12	76	12.1	602	1 PGR1_MOUSE	P22437 mus musculu
13	75.5	12.1	822	1 FGR1_HUMAN	P11362 homo sapien
14	75.5	12.1	822	1 FGR1_MOUSE	P16092 mus musculu
15	75	12.0	359	1 LACH_DROME	Q24372 drosophila
16	74.5	11.9	819	1 FGR1_CHICK	P21804 gallus gall
17	74	11.8	824	1 MYR1_HUMAN	Q9UDY8 homo sapien
18	73.5	11.7	1666	1 MYM1_MOUSE	O65233 mus musculu
19	73	11.7	1447	1 DCC_MOUSE	P70211 mus musculu
20	73	11.6	3038	1 TRIO_HUMAN	O75962 homo sapien
21	72.5	11.6	806	1 CERK2_CHICK	P14660 gallus gall
22	72	11.5	248	1 MYPO_HUMAN	P25189 homo sapien
23	72	11.5	333	1 AMAL_DROME	P13364 drosophila
24	72	11.5	873	1 FAS2_DROME	P34082 drosophila
25	72	11.5	987	1 K6P1_CANAL	O94201 candida alb
26	72	11.5	1070	1 PKR7_HUMAN	O13308 homo sapien
27	71.5	11.4	602	1 PGR1_RAT	O63921 rattus norv
28	71	11.3	654	1 BRR2_HUMAN	O01742 homo sapien
29	70.5	11.3	847	1 CD22_HUMAN	P20273 homo sapien
30	70.5	11.3	887	1 UFO_HUMAN	P30530 homo sapien
31	70.5	11.3	1897	1 P1P5_HUMAN	P10586 homo sapien
32	70	11.2	365	1 CXAR_HUMAN	P78310 homo sapien
33	70	11.2	1447	1 DCC_HUMAN	P43146 homo sapien

ALIGNMENTS

```

RESULT 1
ID      SN_HUMAN          STANDARD:      PRT: 1709 AA.
AC      Q98222; Q9H1H6; Q9H1H7; Q9H7L7; Q9GZ55;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Sialoadhesin precursor (Sialic acid binding Ig-like lectin-1) (Siglec-
DE      1) (CD169 antigen).
CN      SN.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RT      Hartnell A., Steel J., Turley H., Jones M., Jackson D., Crocker P.R.;
RT      "Characterisation of human sialoadhesin (Siglec-1), a sialic acid
RT      binding receptor expressed on resident and inflammatory macrophage
RT      populations.";
RT      Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX      MEDLINE=21638749; PubMed=11780052;
RA      Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA      Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA      Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA      Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA      Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA      Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA      Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA      Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA      Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA      Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA      Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA      Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA      Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA      Lehasialho M.H., Liversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA      Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurtry A.A.,
RA      Milne S.A., Misty D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA      Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA      Phillimore B.J.C.T., Prithalingam S.R., Plumb R.W., Ramsay H.,
RA      Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA      Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Stilson J.E.,
RA      Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA      Tracey A., Triomans A.C., Vaudin M., Wall M., Wallis J.M.,
RA      Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA      Wilming L., Way P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA      Rogers J.;
RA      "The DNA sequence and comparative analysis of human chromosome 20.";
RA      Nature 414:865-871(2001).
RN      [3]
RP      SEQUENCE OF 733-1709 FROM N.A. (ISOFORMS 1 AND 2).
RC      TISSUE=Spleen;
RC      Ohara O., Nagase T., Kikuno R., Okumura K.;
RA      "The nucleotide sequence of a long cDNA clone isolated from human

```


[illegible]

Query Match	Best Local Similarity	Score	83.5%	DB 1	Length	1694		
Matches	37	Conservative	17	Mismatches	42	Indels	75	Gaps
Qy	12	LSPEDQVLPALGSSALNCTAMVYSGPHGSLTPOVLKXGLGLGGH-----	61					
Db	246	LSSSGRIILP---GDPTLTLCR---VNSSYPAVSAVQARQGVMLGTVGHRLRFSAAWN	239					
Oy	62	-----YSLH-----	67					
Db	300	DSGATTCQATNDKGLSSVSPLSLHRYMAEYKKNPAGPVLNEYTYTLTLCSTPKKAPDELR	359					
Oy	68	SWKAN--LSEVLYSVLVGVNVTSTEVYGAFTCSIONISSFSFTLQRAQPT	117					

```

Db      360 SWYNNHLLDHAASHLHPAVVRADTGTFYFCEYONAGS-----ERSPLS 406
RESULT 4
LU_HUMAN ID LU_HUMAN STANDARD; PRT: 628 AA.
AC P50895;
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Lutheran blood group glycoprotein precursor (B-CAM cell surface
DE glycoprotein) (Auburger B antigen) (FB/6253 antigen).
OS Lu OR BCAM OR MSX19.
OC Homo sapiens (human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NC NCLTaxID=9606;
[1]
PC SEQUENCE FROM N.A., AND SEQUENCE OF 32-67 AND 182-203.
RP TISSUE=Placenta;
RX MEDLINE=95296337; Pubmed=7777537;
RX Parsons S.F., Mallinson G., Holmes C.H., Houlihan J.M., Simpson K.L.,
RX Mabry W.J., Spurr N.R., Warne D., Barclay A.N., Anstee D.O.;
RA "The Lutheran blood group glycoprotein, another member of the
RA immunoglobulin superfamily, is widely expressed in human tissues and
RA is developmentally regulated in human liver."
Proc. Natl. Acad. Sci. U.S.A. 92:5496-5500(1995).
[2]
RN RN
RS SEQUENCE OF 1-588 FROM N.A.
RX MEDLINE=95042297; Pubmed=7954395;
RX Campbell I.G., Foulkes W.D., Senger G., Trowsdale J.,
RX Garin-Chesa P., Rettig W.J.;
RA "Molecular cloning of the B-CAM cell surface glycoprotein of
RA epithelial cancers: a novel member of the immunoglobulin
RA superfamily."
Cancer Res. 54:5761-5765(1994).
RL RL
CC -1 FUNCTION: PROBABLE RECEPTOR. MAY MEDIATE INTERCELLULAR SIGNALING.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 TISSUE SPECIFICITY: WIDE-TISSUE DISTRIBUTION (HIGHEST IN THE
CC PANCREAS AND VERY LOW IN BRAIN). CLOSELY ASSOCIATED WITH THE BASAL
CC LAYER OF CELLS IN EPITHELIA AND THE ENDOTHELIUM OF BLOOD VESSEL
CC WALLS.
CC -1 DEVELOPMENTAL STAGE: IS UNDER DEVELOPMENTAL CONTROL IN LIVER AND
CC MAY ALSO BE REGULATED DURING DIFFERENTIATION IN OTHER TISSUES.
CC UPREGULATED FOLLOWING MALIGNANT TRANSFORMATION IN SOME CELL TYPES.
CC -1 POLYMORPHISM: LU IS RESPONSIBLE FOR THE LUTHERAN BLOOD GROUP
CC SYSTEM.
CC -1 SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1 SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1 SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X83425; CAA58449.1; -.
DR EMBL: X80026; CAAS6327.1; -.
DR Genew: HGNC:6722; LU.
DR MIM: 111200; -.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig_5.
DR SMART: SM00410; Ig_Like; 2.
DR SHAT: SM00408; IgC2; 2.
DR Receptor: Immunoglobulin domain; Glycoprotein; Transmembrane; Signal;
DM Repeat: Blood group antigen.
LT SIGNAL 31

```

```
FT CHAIN 32 628 LUTHERAN BLOOD GROUP GLYCOPROTEIN.
FT DOMAIN 32 547 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 548 568 POTENTIAL.
FT DOMAIN 569 628 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 46 132 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 165 244 IG-LIKE V-TYPE DOMAIN 2.
FT DOMAIN 284 344 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 377 431 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 466 529 IG-LIKE C2-TYPE DOMAIN 3.
FT DISULFID 53 125 PROBABLE.
FT DISULFID 172 237 PROBABLE.
FT DISULFID 291 337 PROBABLE.
FT DISULFID 384 424 PROBABLE.
FT DISULFID 473 522 PROBABLE.
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 377 377 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 225 226 RL -> PC (IN REF. 2).
FT CONFLICT 355 356 EL -> DV (IN REF. 2).
SO SEQUENCE 628 AA: 67374 MW: C89B0A4835492B1E CRC64:

Query Match 13.1%; Score 82; DB 1; Length 628;
Best Local Similarity 36.2%; Pred. No. 0.95;
Matches 29; Conservative 12; Mismatches 33; Indels 6; Gaps 3;

OY 10 DFLSSENOVLRPALGSSVALNCTAMVSGPHCSLPYQWLKDLPLIGHSLSHEXSW 69
Db 362 DPLEISEKGVLSLPLNSSAVNCS--VHG--LPPPALRWTKDPSLDGPMWLSLSITF 416
OY 70 VKANLEVLVSVLGAVNTS 89
Db 417 -DSNCTVYCEASLPYVPLVS 435

RESULT 5
FGR4_MOUSE STANDARD: PRT: 808 AA.
ID FGR4_MOUSE
AC 003142:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibroblast growth factor receptor 4 precursor (EC 2.7.1.112) (FGR-4)
DE (Protein-tyrosine kinase receptor MPK-11).
CN FGR4 OR FGR-4 OR MPK-11.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
NCBI_TaxID=10090;
OX (1)
RP SEQUENCE FROM N.A.
RC STRAIN-CD-1; TISSUE-Fetal cerebellum;
RX MEDLINE-92146274; Pubmed-1723680;
RA Stark K.L., McMahon J., McMahon A.P.;
RT "FGR-4, a new member of the fibroblast growth factor receptor
RT family, expressed in the definitive endoderm and skeletal muscle
RT lineages of the mouse."
RL Development 113:641-651(1991).
RN (2)
RP SEQUENCE OF 620-676 FROM N.A.
RC STRAIN-C57BL/6; TISSUE-Embryonic brain;
RX MEDLINE-93096484; Pubmed-1281307;
RA Glazard-Hedenstreit P., Nieto M.A., Frain M., Mattei M.-G.,
RA Chesnier A., Wilkinson D.G., Charney P.;
RT "An Epi-related receptor protein tyrosine kinase gene segmentally
RT expressed in the developing mouse hindbrain."
RL Oncogene 7:2499-2506(1992).
CC -1- FUNCTION: PUTATIVE RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC MAY BE INVOLVED IN THE DEVELOPMENT OF SKELETAL MUSCLE CELL
CC LINEAGES.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
```

```
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING GUT ENDODERM,
CC IN MYOTOMALLY DERIVED SKELETAL MUSCLE, THE ADRENAL CORTEX,
CC KIDNEY AND CONDENSING CARTILAGE. PRESENT IN ADULT LIVER,
CC LUNG AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X59927; CAA42551.1; -.
CC EMBL: X57236; CAA40512.1; -.
CC HSSP: P11362; IFCG.
CC MGD: MGI:95525; Fgfr4.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003598; Ig_C2.
CC InterPro: IPR01245; Tyr_pkinase.
CC Pfam: PF00047; Ig_3.
CC Pfam: PF00069; pkinase.
CC PRINTS: PR00109; TYRKINASE.
CC PRODOM: PD000001; Euk_pkinase. 1.
CC SMART: SM00408; IGC2_3.
CC SMART: SM00219; TyKc_1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP_FALSE_NEG.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR_1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
CC Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
CC Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
CC Repeat; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 808 FIBROBLAST GROWTH FACTOR RECEPTOR 4.
FT DOMAIN 19 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 POTENTIAL.
FT DOMAIN 388 808 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 53 98 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 169 221 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 268 330 IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 464 761 PROTEIN KINASE.
FT NP_BIND 470 478 ATP (BY SIMILARITY).
FT BINDING 509 509 ATP (BY SIMILARITY).
FT ACT_SITE 618 618 BY SIMILARITY.
FT MOD_RES 649 649 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 53 98 POTENTIAL.
FT DISULFID 169 221 POTENTIAL.
FT DISULFID 268 330 POTENTIAL.
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 808 AA: 89775 MW: E3F30BE5334E86 CRC64:

Query Match 12.9%; Score 80.5; DB 1; Length 808;
Best Local Similarity 30.8%; Pred. No. 1.8;
Matches 28; Conservative 11; Mismatches 37; Indels 15; Gaps 4;

OY 24 LGSSVALNCTAMVSGPHCSLPYQWLK---DGLPLDIGHSLSHEXSWKANISEVLV 79
Db 260 VGSVDLCKKYSDAQPH-----IQWLKHVVINGSSFGADGPPYQVQLKTTDINISEV-- 312
OY 80 SSVLCVNTSTREVCAPFCISQNN---ISSS 107
Db 313 -QVLYLRNVSADGEVYCLAGNIGSLSTYS 342
```

RESULT 6
DSCA_HUMAN STANDARD: PRT: 2012 AA.
AC 060469; 060468; Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Down syndrome cell adhesion molecule precursor (CHD2).
GN DSCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=98087574; PubMed=9426258;
RA Yamakawa K., Huot Y.-K., Haendel M.A., Hubert R., Chen X.-N.,
RA Lyons G.E., Kornberg J.R.;
RT Down syndrome region of the immunoglobulin superfamily maps in a
PE nervous system. Hum. Mol. Genet. 7:227-237(1998).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=20384934; PubMed=10925149;
RA Agarwala K.L., Nakamura S., Tsutsumi Y., Yamakawa K.;
RT "Down syndrome cell adhesion molecule DSCAM mediates homophilic
RT intercellular adhesion.";
RL Brain Res. Mol. Brain Res. 79:118-126(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Matsubae H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Toriki Y., Choi D.-K., Soda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Rosenthal K., Rump A., Schilpabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kuch J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Anttonen S.E.,
RA Minoshima S., Shimizu N., Nordstieck G., Horstischer K., Brandt P.,
RA Scharte M., Schen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramer J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- FUNCTION: CELL ADHESION MOLECULE THAT CAN MEDIATE CATION-
CC INDEPENDENT HOMOPHILIC BINDING ACTIVITY. COULD BE INVOLVED IN
CC NERVOUS SYSTEM DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PROBABLE). THE
CC SHORT ISOFORM MAY BE SECRETED.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM/CHD2-52 (SHOWN HERE)
CC AND A SHORT FORM/CHD2-42; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 10 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF023450; AAC17967.1; -
CC EMBL: AF023449; AAC17966.1; -
CC EMBL: AF217525; AAF27525.1; -
CC EMBL: AL163283; CAB90464.1; -
CC EMBL: AL163282; CAB90436.1; -
CC EMBL: AL163281; CAB90444.1; -

DR Genew: HGNC:3039; DSCAM.
DR MIM: 602525;
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FNIII_repeat.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003600; IG_Like.
DR Pfam: PF00041; fn3; 6.
DR Pfam: PF00047; fn3; 10.
DR PRINTS: PR00014; ENTRYPELII.
DR SMART: SM00060; FN3; 6.
DR SMART: SM00410; IG_Like; 2.
DR SMART: SM00408; IGc2; 7.
KW Immunoglobulin domain; Glycoprotein; Signal; Cell adhesion; Repeat;
KW Transmembrane; Alternative splicing
FT SIGNAL 1
FT CHAIN 17
FT DOMAIN 18 2012
FT TRANSMEM 18 1593
FT DOMAIN 1596 1616
FT DOMAIN 1617 2012
FT DOMAIN 39 109
FT DOMAIN 138 204
FT DOMAIN 239 300
FT DOMAIN 328 392
FT DOMAIN 421 491
FT DOMAIN 491 582
FT DOMAIN 518 676
FT DOMAIN 610 773
FT DOMAIN 704 872
FT DOMAIN 802 972
FT DOMAIN 984 1076
FT DOMAIN 1088 1177
FT DOMAIN 1189 1323
FT DOMAIN 1360 1366
FT DOMAIN 1380 1463
FT DOMAIN 1477 1562
FT DISULFID 46 102
FT DISULFID 145 197
FT DISULFID 246 293
FT DISULFID 335 385
FT DISULFID 428 484
FT DISULFID 525 575
FT DISULFID 617 669
FT DISULFID 711 766
FT DISULFID 809 859
FT DISULFID 1307 1359
FT CARBOHYD 28 78
FT CARBOHYD 78 28
FT CARBOHYD 470 470
FT CARBOHYD 487 487
FT CARBOHYD 512 512
FT CARBOHYD 556 556
FT CARBOHYD 658 658
FT CARBOHYD 666 666
FT CARBOHYD 710 710
FT CARBOHYD 710 710
FT CARBOHYD 748 748
FT CARBOHYD 795 795
FT CARBOHYD 924 924
FT CARBOHYD 1142 1142
FT CARBOHYD 1160 1160
FT CARBOHYD 1250 1250
FT CARBOHYD 1271 1271
FT CARBOHYD 1341 1341
FT CARBOHYD 1488 1488
FT CARBOHYD 1562 1562
FT VARSPPLIC 1572
FT VARSPPLIC 1571
FT VARSPPLIC 1572
FT CONFLICT 1893
SQ SEQUENCE 2012 AA: 222259 MW: 0E33CEB781A0B334 CXC64;
MISSING (IN SHORT ISOFORM).
HRRGDLHPPYIPMBLLARRGGPTSRDLSGQAALEPK
SRTRKRPVLEPIPMELASASSTREGCSMPGAVATLPR
EGATLGCAAKHSSQBSLDSRCHIKGNPYAKSYTLV ->
IGVTSYICLHLEMTYC (IN REF. 1).
MW: 0E33CEB781A0B334 CXC64;

Query Match 12.8%; Score 80; DB 1; Length 2012;
 Best Local Similarity 26.3%; Pred. No. 5.6;
 Matches 26; Conservative 19; Mismatches 38; Indels 16; Gaps 5;

QY 6 DRAPPPLSPEDQVLRPALGSSVALNCTAMVVSGRPCSLPSVQWIKDGLPLIGIHSYSH 65
 DB 404 DGTPTIISAFSEKVVSPA--EPVSLMCN---VKG--TLPPTITWLDLDDPLIKGSHRIS 456
 QY 66 EYSWKANSEVLVSSVGVNTSTREYV--GAFCSION 102
 DB 457 QMITSEGNVSTYL-----NISSSQVRDGGVYRGTANN 488

RESULT 7
 NPHN_MOUSE
 ID NPHN_MOUSE STANDARD; PRT: 1242 AA.
 AC 090257;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nephrin precursor (renal glomerulus-specific cell adhesion receptor).
 GN NPHS1 OR NPHN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
 RX MEDLINE=99436348; PubMed=10504499;
 RA Holman L.B., St John P.L., Kovari I.A., Verma R., Holtfoefer H.,
 RA Abrahamson D.R.;
 RT "Nephrin localizes to the slit pore of the glomerular epithelial
 cell.";
 RT Kidney Int. 56:1481-1491(1999).
 RL [2]
 RN [2]
 RP INTERACTION WITH CD2AP.
 RX MEDLINE=21590051; PubMed=11733379;
 RA Shih N.Y., Li J., Cotran R., Mundel P., Miner J.H., Shaw A.S.;
 RT "CDAP localizes to the slit diaphragm and binds to nephrin via a
 novel C-terminal domain.";
 RT Am. J. Pathol. 159:2303-2308(2001).
 RN [3]
 RP INTERACTION WITH CD2AP AND NPHS2.
 RX MEDLINE=21590460; PubMed=11733557;
 RA Schatz K., Simons M., Reiser J., Saleem M.A., Paul C., Kriz W.,
 RA Shaw A.S., Holzman L.B., Mundel P.;
 RT "Podocin, a raft-associated component of the glomerular slit
 diaphragm, interacts with CD2AP and nephrin.";
 RT J. Clin. Invest. 108:1621-1629(2001).
 CC -1- FUNCTION: Seems to play a role in the development or function of
 the kidney glomerular filtration barrier. May anchor the podocyte
 slit diaphragm to the actin cytoskeleton.
 CC -1- SUBUNIT: Interacts with podocin/NPHS2 and with CD2AP C-terminal
 domain.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential). Located
 at podocyte slit diaphragm between podocyte foot processes.
 CC -1- TISSUE SPECIFICITY: Expressed in kidney glomeruli.
 CC -1- PTM: Phosphorylated on tyrosine residues (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 8 IMMUNOGLOBULIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL: AF168466; AAF03368.1; -
 CC MGI: MGI:1859637; Nphs1.
 DR InterPro: IPR003961; FN_III.

DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR00399; TTP_enzyme.
 DR Pfam: PF00041; fn3; 1.
 DR Pfam: PF00047; Ig; 8.
 DR SMART: SM00060; FN3; 1.
 DR SMART: SM00408; FN3; 1.
 DR SMART: SM00410; Ig_Like; 3.
 KW Cell adhesion; transmembrane; Signal; Glycoprotein;
 KW Immunoglobulin domain; Repeat; Phosphorylation.
 FT SIGNAL 1 22
 FT CHAIN 23 1242
 FT DOMAIN 23 1064
 FT TRANSMEM 1065 1086
 FT DOMAIN 1087 1242
 FT DOMAIN 46 118
 FT DOMAIN 153 224
 FT DOMAIN 258 324
 FT DOMAIN 354 424
 FT DOMAIN 458 535
 FT DOMAIN 560 630
 FT DOMAIN 754 823
 FT DOMAIN 856 927
 FT DOMAIN 941 1025
 FT DISULFID 53 111
 FT DISULFID 160 217
 FT DISULFID 265 317
 FT DISULFID 361 417
 FT DISULFID 465 528
 FT DISULFID 567 623
 FT DISULFID 761 816
 FT DISULFID 863 920
 FT CARBOHYD 40 40
 FT CARBOHYD 356 356
 FT CARBOHYD 401 401
 FT CARBOHYD 547 547
 FT CARBOHYD 553 553
 FT CARBOHYD 564 564
 FT CARBOHYD 577 577
 FT CARBOHYD 680 680
 FT CARBOHYD 708 708
 FT CARBOHYD 908 908
 FT SEQUENCE 1242 AA; 134890 MW; 02DB2180BF145092 CRC64;

Query Match 12.7%; Score 79.5; DB 1; Length 1242;
 Best Local Similarity 29.5%; Pred. No. 3.7;
 Matches 26; Conservative 16; Mismatches 29; Indels 17; Gaps 5;

QY 26 SSVALNCTAMVVSGRPCSLPSVQ--WIKDGLPLGI-GGHSYSHSEYWKANSEVLVSSV 82
 DB 857 SSATLHCRA-----RGVNDIFWTWKGVLDDLPDPRTEHKYH-----QGVVHSSL 903
 QY 83 LGV-NVTSREYVGAFTCSIONISFSSET 109
 DB 904 LRTIANVSAQDYALFKCTATNALGSDHT 931

RESULT 8
 NPHN_MOUSE
 ID NPHN_MOUSE STANDARD; PRT: 1234 AA.
 AC 090404; 090404;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nephrin precursor (renal glomerulus-specific cell adhesion receptor).
 GN NPHS1 OR NPHN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=101116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).

FT	VARIANT	743	743	R -> C (IN CNF).	/FtId=VAR_013048.
FT	VARIANT <td>802<td>802</td><td>R -> W (IN CNF).<th>/FtId=VAR_013049.</th></td></td>	802 <td>802</td> <td>R -> W (IN CNF).<th>/FtId=VAR_013049.</th></td>	802	R -> W (IN CNF). <th>/FtId=VAR_013049.</th>	/FtId=VAR_013049.
FT	VARIANT <td>802<td>802</td><td>R -> P (IN CNF).<th>/FtId=VAR_013050.</th></td></td>	802 <td>802</td> <td>R -> P (IN CNF).<th>/FtId=VAR_013050.</th></td>	802	R -> P (IN CNF). <th>/FtId=VAR_013050.</th>	/FtId=VAR_013050.
FT	VARIANT <td>806<td>806</td><td>A -> D (IN CNF).<th>/FtId=VAR_013051.</th></td></td>	806 <td>806</td> <td>A -> D (IN CNF).<th>/FtId=VAR_013051.</th></td>	806	A -> D (IN CNF). <th>/FtId=VAR_013051.</th>	/FtId=VAR_013051.
FT	VARIANT <td>819<td>819</td><td>D -> V (IN CNF).<th>/FtId=VAR_013052.</th></td></td>	819 <td>819</td> <td>D -> V (IN CNF).<th>/FtId=VAR_013052.</th></td>	819	D -> V (IN CNF). <th>/FtId=VAR_013052.</th>	/FtId=VAR_013052.
FT	VARIANT <td>831<td>831</td><td>R -> C (IN CNF).<th>/FtId=VAR_013053.</th></td></td>	831 <td>831</td> <td>R -> C (IN CNF).<th>/FtId=VAR_013053.</th></td>	831	R -> C (IN CNF). <th>/FtId=VAR_013053.</th>	/FtId=VAR_013053.
FT	VARIANT <td>831<td>831</td><td>R -> C (IN CNF).<th>/FtId=VAR_013053.</th></td></td>	831 <td>831</td> <td>R -> C (IN CNF).<th>/FtId=VAR_013053.</th></td>	831	R -> C (IN CNF). <th>/FtId=VAR_013053.</th>	/FtId=VAR_013053.
Query Match					
Best Local Similarity		12.2%	Score 76.5;	DB 1;	Length 1241;
Matches	29;	Conservative	16;	Mismatches	34;
				Indels	23;
				Gaps	6;
OY	26	SVALINCTAWVYSGPHCLPSV--QWLKDGLPLGI-GGHYSLSHEYSWKANLSEYLVSSV	82		
DB	857	SSATLHCHRA-----RQVPIVFTWNTGNGVPLDQDRVTEHYHOGGVH-----SSL	903		
OY	83	IGV-NVTSNEYGAFCTSGIONISFSSFT-----LQRAGPS	117		
DB	904	LTIANVSAADYALFTCTATNALGSDQINIQLVSTISRPDPS	945		
RESULT 12					
PGH1_MOUSE					
ID	PGH1_MOUSE	STANDARD:	PRT:	602 AA.	
AC	P22437.				
DT	01-AUG-1991 (Rel. 19, Created)				
DR	01-AUG-1991 (Rel. 19, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Prostaglandin G/H synthase 1 precursor (EC 1.14.99.1) (Cyclooxygenase				
DE	-1) (COX-1) (Prostaglandin-endoperoxide synthase 1) (Prostaglandin H2				
DE	synthase 1) (PGH synthase 1) (PGHS-1) (PHS 1).				
GN	PTGS1 OR COX1 OR COX-1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90203007; PubMed=2108169;				
RA	Dewitt D.L., El-Hariri E.A., Kraemer S.A., Andrews M.J., Yao E.F.,				
RA	Armstrong R.L., Smith W.L.				
RT	"The aspirin and heme-binding sites of ovine and murine prostaglandin				
RT	endoperoxide synthases.";				
RL	J. Biol. Chem. 265:5192-5198(1990).				
CC	-1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN REGULATING OR PROMOTING				
CC	CELL PROLIFERATION IN SOME NORMAL AND NEOPLASTICALLY TRANSFORMED				
CC	CELLS.				
CC	-1- CATALYTIC ACTIVITY: Arachidonate + AH(2) + 2 O(2) = prostaglandin				
CC	H2 + A + H(2O).				
CC	-1- PATHWAY: FIRST STEP IN THE FORMATION OF PROSTAGLANDINS AND				
CC	THROMBOXANES.				
CC	-1- SUBUNIT: HOMODIMER.				
CC	-1- SUBCELLULAR LOCATION: Membrane-associated. Microsomal membrane.				
CC	-1- MISCELLANEOUS: THIS ENZYME ACTS BOTH AS A DIOXYGENASE AND AS A				
CC	PEROXIDASE.				
CC	-1- MISCELLANEOUS: THIS ENZYME IS THE TARGET OF NONSTEROIDAL				
CC	ANTI-INFLAMMATORY DRUGS SUCH AS ASPIRIN.				
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.				
CC	-1- SIMILARITY: BELONGS TO THE PROSTAGLANDIN G/H SYNTHASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				

EMBL: M34141; AAA39913.1; -
DR PIR: A35564; A35564.
DR HSSP: P05979; 1DIY..
DR MGD: MGI:97797; PTGSL.
DR InterPro: IPR002007; Anln_peroxidase.
DR InterPro: IPR00361; EGF-like.
DR Pfam: PF03098; EGF_1.
DR Pfam: PF03098; An_peroxidase; 1.
DR PRINTS: SM00181; EGF; 1.
DR SMART: SM00457; ANPEROXIDASE.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.
DR Oxidoreductase; Dioxigenase; Peroxidase; Glycoprotein; Acetylation;
KW Prostaglandin biosynthesis; Heme; Iron; Signal; Membrane;
KW EGF-like domain.
FT SIGNAL 1 26
FT CHAIN 27 602 PROSTAGLANDIN G/H SYNTHASE 1.
FT DOMAIN 34 72 EGF-LIKE.
FT ACT_SITE 209 209 DISTAL HISTIDINE (BY SIMILARITY).
FT ACT_SITE 387 387 CYCLOOXYGENASE (BY SIMILARITY).
FT BINDING 390 390 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT MOD_RES 532 532 ASPIRIN-ACETYLATED SERINE.
FT DISULFID 43 49 BY SIMILARITY.
FT DISULFID 61 59 BY SIMILARITY.
FT DISULFID 61 71 BY SIMILARITY.
FT DISULFID 39 71 BY SIMILARITY.
FT DISULFID 571 577 BY SIMILARITY.
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 602 AA; 634C0E602045C3A0 CRC64;
Query Match 12.1%; Score 76; DB 1; Length 602;
Best Local Similarity 27.4%; Pred. No. 3.6;
Matches 34; Conservative 15; Mismatches 59; Indels 16; Gaps 6;
OY 2 PGVCDRADPLSPSEDOVLEPALGSSVAL-----NCTAVVSGPCISLPSV-QMLK 51
DB 22 PSYLLADPVPSPVNPCCYPCQNGCVRFGLDNYQCDCOTRNGSGPACTIPEIMTWLR 81
OY 52 DGL-PLGIGGHVSL-HEY-SWVKAN---LSEVYSSVLGVNTSTVEGAFCSIONISF 105
DB 82 NSLRPSSFTFLHTHCYWMIEFVNATFIREVLRVLVYRSRLTSPPTYSANHYISW 141
OY 106 SFTT 109
DB 142 ESFS 145
UT 13
ID HUMAN STANDARD: PRT; 822 AA.
AC P11362; P17049;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Basic fibroblast growth factor receptor 1 precursor (BC 2.7.1.112)
DE (EGFR-1) (bFGF-R) (Fms-like tyrosine kinase-2) (c-fgr).
GN FGFR1 OR FLG OR FGFR OR FLR2.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RP
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RX MEDLINE=90245600; PubMed=2159626;
RA Isaac H. A., Bergonzoni L., Sarmientos P.;
RT "Complete sequence of a human receptor for acidic and basic
RL fibroblast growth factors.";
RL Nucleic Acids Res. 18:1906-1906(1990).
RP
RP SEQUENCE FROM N.A.

RC TISSUE-Neonatal brain stem;
RX MEDLINE=90360977; PubMed=1697263;
RA Dionne C.A., Crumley G.R., Ballot F., Kaplow J.M., Searfoss G.,
RA Ruta M., Burgess W.H., Jey M., Schlössinger J., Terada M.,
RT "Cloning and expression of two distinct high-affinity receptors
RT cross-reacting with acidic and basic fibroblast growth factors.";
RL EMBO J. 9:2685-2692(1990).
[3]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=92282615; PubMed=1317750; Sakamoto H., Morita T.,
RA Hattori Y., Odagiri H., Katoh O., Sugimura T., Terada M.,
RA Shimocho K., Tobinal K., Sugimura T., Terada M.,
RT "K-sam-related gene, N-sam, encodes fibroblast growth factor receptor
RT and is expressed in T-lymphocytic tumors.";
RL Cancer Res. 52:3367-3371(1992).
[4]
RN
RP SEQUENCE FROM N.A.
RC TISSUE-Liver;
RX MEDLINE=91126480; PubMed=1846977;
RA Hou J., Kan M., McKeenan K., McBride G., Adams P., McKeenan W.L.;
RT "Fibroblast growth factor receptors from liver vary in three
RT structural domains.";
RL Science 251:665-668(1991).
[5]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=92118399; PubMed=1662973;
RA Kiefer M.C., Baird A., George-Nascimento C., Nguyen T., Mason O.B.,
RA Boley L.J., Valenzuela P., Barr P.J.;
RT "Molecular cloning of a human basic fibroblast growth factor receptor
RT cDNA and expression of a biologically active extracellular domain in
RL a baculovirus system.";
RN Growth factors 5:115-127(1991).
[6]
RN
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RX MEDLINE=90290512; PubMed=2162671;
RA Itoh N., Terachi T., Ohta M., Seo M.K.;
RT "The complete amino acid sequence of the shorter form of human basic
RT fibroblast growth factor receptor deduced from its cDNA.";
RL Biochem. Biophys. Res. Commun. 169:680-685(1990).
[7]
RN
RP SEQUENCE OF 201-822 FROM N.A.
RA Ruta M., Hawk R., Rizza G., Drohan W., Zabejshansky M., Laureys G.,
RA Barton D.E., Francke U., Schlössinger J., Givol D.;
RT "A novel protein tyrosine kinase gene whose expression is modulated
RT during endothelial cell differentiation.";
RL Oncogene 3:9-15(1988).
[8]
RN
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=90355989; PubMed=2167437;
RA Johnson D.E., Lee P.L., Lu J., Williams L.T.;
RT "Diverse forms of a receptor for acidic and basic fibroblast growth
RT factors.";
RL Mol. Cell. Biol. 10:4728-4736(1990).
[9]
RN
RP ALTERNATIVE SPLICING.
RX MEDLINE=91144499; PubMed=1847500;
RA Gunkel S.J., Link D.C., Katamine S., Lacal P., Miki T., Ley T.J.,
RA Robbins K.C.;
RT "A novel c-fgr exon utilized in Epstein-Barr virus-infected B
RT lymphocytes but not in normal monocytes.";
RL Mol. Cell. Biol. 11:1500-1507(1991).
[10]
RN
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE-Lung;
RX MEDLINE=91319400; PubMed=1650441;
RA Eisenmann A., Ahn J.A., Graziani G., Tronick S.R., Ron D.;
RT "Alternative splicing generates at least five different isoforms of
RL the human basic-FGF receptor.";
RL Oncogene 6:1195-1202(1991).
[11]
RN
RP SEQUENCE FROM N.A.
RC Wennstroem S., Sandstroem C., Claesson-Welsh L.;

Submitted (JUL-1990) to the EMBL/Genbank/DBJ databases.

[12] MUTAGENESIS OF TYR-766.
 RA MEDLINE-92357144; PubMed-1379697;
 RA Peters K.G., Marie J., Wilson E., Ives H.E., Escobedo J.,
 RA del Rosario M., Mirda D., Williams L.T.;
 RT "Point mutation of an FGF receptor abolishes phosphatidylinositol
 turnover and Ca²⁺ flux but not mitogenesis.";
 RL Nature 358:678-681(1992).

[13] MUTAGENESIS OF TYR-766.
 RA MEDLINE-92357145; PubMed-1379698;
 RA Mohammad M., Dionne C.A., Li W., Lin N., Spivak T., Honegger A.M.,
 RA Jave M., Schlessinger J.;
 RT "Point mutation in FGF receptor eliminates phosphatidylinositol
 hydrolysis without affecting mitogenesis.";
 RL Nature 358:681-684(1992).

[14] X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 464-762.
 RA MEDLINE-96361355; PubMed-8752212;
 RA Mohammad M., Schlessinger J., Hubbard S.R.;
 RT "Structure of the FGF receptor tyrosine kinase domain reveals a novel
 autoinhibitory mechanism.";
 RL Cell 86:577-587(1996).

[15] X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 464-762.
 RA MEDLINE-97284786; PubMed-9139660;
 RA Mohammad M., McMahon G., Sun L., Tang C., Hirth P., Yeh B.K.,
 RA Hubbard S.R., Schlessinger J.;
 RT "Structures of the tyrosine kinase domain of fibroblast growth factor
 receptor in complex with inhibitors.";
 RL Science 276:955-960(1997).

[16] VARIANT PEIFFER SYNDROME ARG-252.
 RA MEDLINE-95179173; PubMed-7874169;
 RA Muenke L., Schei U., Hehr A., Robin N.R., Losken H.W., Schinzel A.,
 RA Puley L.J., Rutland P., Reardon W., Malcolm S., Winter R.M.;
 RT "A common mutation in the fibroblast growth factor receptor 1 gene in
 Pfeiffer syndrome.";
 RL Nat. Genet. 8:269-274(1994).

-1 FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR. A SHORTER
 FORM OF THE RECEPTOR COULD BE A RECEPTOR FOR ACIDIC FGF (AFGF).
 CC -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 tyrosine phosphate.
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 ALTERNATIVE PRODUCTS: MANY FORMS OF FGFR1 ARE PRODUCED BY
 ALTERNATIVE SPLICING. THE FORM SHOWN HERE IS KNOWN AS ALPHA-A1.
 CC -1 DISEASE: DEFECTS IN FGFR1 ARE ONE OF THE CAUSES OF PEIFFER
 SYNDROME (FS) (ALSO KNOWN AS ACROCEPHALOSYNDACTYLY TYPE V; ACS5);
 CHARACTERIZED BY CRANIOSYNOSTOSIS (PREMATURE FUSION OF THE SKULL
 SUTURES) WITH DEVIATION AND ENLARGEMENT OF THE THUMBS AND GREAT
 TOES, BRACHYMESEPHALANGY, WITH PHALANGAL ANKYLOSIS AND A VARYING
 DEGREE OF SOFT TISSUE SYNDACTYLY.
 CC -1 DISEASE: INVOLVED IN A T(8;13)(p12;q12) CHROMOSOMAL TRANSLOCATION
 WHICH INVOLVES FGFR1 AND ZNF198. The resulting transcript is a
 possible candidate for stem cell leukemia lymphoma syndrome/SCLL.
 CC -1 SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
 FAMILY.
 CC -1 SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1 DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
 WWW="http://www.infobiogen.fr/services/chronocancer/Genes/FGFR1113.html".

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; X51803; CAA36101.1; -
 DR EMBL; X52833; CAA37015.1; -
 DR EMBL; X66945; CAA47375.1; -

DR EMBL; Y00665; CAA68679.1; -
 DR EMBL; M37722; AAA75007.1; -
 DR EMBL; M60485; AAA35840.1; -
 DR EMBL; M63887; AAA35958.1; -
 DR EMBL; M34185; AAA35836.1; -
 DR EMBL; M34186; AAA35837.1; -
 DR EMBL; X57118; CAA40400.1; ALT_TERM.
 DR EMBL; X57119; CAA40401.1; -
 DR EMBL; X57120; CAA40402.1; -
 DR EMBL; X57121; CAA40403.1; -
 DR EMBL; X57122; CAA40404.1; -
 DR EMBL; M34641; AAA35835.1; -
 DR EMBL; A29216; CAA01958.1; -
 DR PIR; S11692; TVHUNG.
 DR PIR; S25420; S25420.
 DR PIR; S26739; S26739.
 DR PDB; 1FGI; 23-JUL-97.
 DR PDB; 1FGI; 08-APR-98.
 DR PDB; 1AGW; 25-MAR-98.
 DR Genew; HGNC:3688; FGFR1.
 DR MIM; 136350; -
 DR MIM; 101600; -
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig_3.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR Prodom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00219; TYRK; 1.

Query Match 12.1%; Score 75.5; DB 1; Length 822;
 Best Local Similarity 31.3%; Pred. No. 5.8;
 Matches 26; Conservative 9; Mismatches 31; Indels 17; Gaps 4;

QY 23 AGGSAVLAICTAWYVSGPRHCSLPVQWMLDGLPLGIGHSLSHEYSWVAVANSEVLYSSV 82
 DB 268 AGGSNVEFMCK--VYSDPQ--PHIQWLK-----HIEVNGSKRGPDNPVQYLRT 313
 QY 83 LGVNVSTVEYGAFTCSIONISF 105
 DB 314 AGVNTTDKEME--VLHKNVSF 333

RESULT 14
 FGFR1_MOUSE STANDARD; PRT; 822 AA.
 AC P16092; Q01736; Q01562;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Basic fibroblast growth factor receptor 1 precursor (BC 2.7.1.112)
 GN (FGFR-1) (bFGF-R) (MFR).
 CN FGFR1 OR FGF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE-90160373; PubMed-1689490;
 RA Reid H.H., Wilks A.F., Bernard O.;
 RT "Two forms of the basic fibroblast growth factor receptor-like mRNA
 are expressed in the developing mouse brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:1596-1600(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Brain;
 RX MEDLINE-90265603; PubMed-2161096;
 RA Saffran A., Avivi A., Orr-Urtreger A., Neufeld G., Lomai P.,

RA Givol D., Yarden Y.;
 RT "The murine lig gene encodes a receptor for fibroblast growth
 RT factor.";
 RT Oncogene 5:635-643(1990).
 RN [3]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RA MEDLINE-91207411, PubMed-1108247.
 RA Koubara H., Kasayama S., Saito H., Matsumoto K., Sato B.;
 RT "Expression cDNA cloning of fibroblast growth factor (FGF) receptor
 RT in mouse breast cancer cells: a variant form in FGF-responsive
 RT transformed cells.";
 RT Biochem. Biophys. Res. Commun. 176:31-37(1991).
 RN [4]
 RP SEQUENCE FROM N.A. (SHORT FORM).
 RX MEDLINE-90272715, PubMed-2161540;
 RA Manukhani A., Moscatelli D., Talarico D., Levytska V., Basillio C.;
 RT "A murine fibroblast growth factor (FGF) receptor expressed in CHO
 RT cells is activated by basic FGF and Kaposi FGF.";
 RT Proc. Natl. Acad. Sci. U.S.A. 87:4378-4382(1990).
 RN [5]
 RP SEQUENCE OF 1-15 FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE-95100926, PubMed-7802632;
 RA Harada T., Saito H., Koubara H., Kurebayashi S., Kasayama S.,
 RA Terakawa N., Kishimoto T., Sato B.;
 RT "Murine fibroblast growth factor receptor 1 gene generates multiple
 RT messenger RNAs containing two open reading frames via alternative
 RT splicing.";
 RT Biochem. Biophys. Res. Commun. 205:1057-1063(1994).
 CC -1- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR. A SHORTER
 CC FORM OF THE RECEPTOR COULD BE A RECEPTOR FOR ACIDIC FGF (AFGF).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS. A LONG FORM (SHOWN HERE), A
 CC VARIANT FORM AND A SHORT FORM ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M28998; AAA37290.1; -;
 DR EMBL: X51893; AAA36175.1; -;
 DR EMBL: M65053; AAA37620.1; -;
 DR EMBL: M33760; AAA37622.1; -;
 DR EMBL: S74765; AAA32845.1; ALF_SEQ.
 DR PIR: A34849; TYMSRG.
 DR HSP: P11882; IFGR.
 DR MGD: MGI:95522; Fgfr1.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR003598; Iq_C2.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00047; Iq; 3.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00408; IGC2; 3.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
 KW Receptor; Glycoprotein; Tyrosine protein kinase; ATP-binding;
 KW transferase; Phosphorylation; Transmembrane; Signal;
 KW Immunoglobulin domain; Repeat; Alternative splicing.
 FT SIGNAL 1 21
 FT POTENTIAL.

FT CHAIN 22 822
 FT DOMAIN 22 376
 FT TRANSMEM 377 397
 FT DOMAIN 398 822
 FT DOMAIN 48 108
 FT DOMAIN 171 237
 FT DOMAIN 270 348
 FT DOMAIN 478 767
 FT NP_BIND 484 492
 FT BINDING 514 514
 FT ACT_SITE 623 623
 FT MOD_RES 654 654
 FT DISULFID 55 101
 FT DISULFID 178 230
 FT CARBOHYD 277 341
 FT CARBOHYD 117 117
 FT CARBOHYD 227 227
 FT CARBOHYD 240 240
 FT CARBOHYD 264 264
 FT CARBOHYD 296 296
 FT CARBOHYD 317 317
 FT CARBOHYD 330 330
 FT VARSPPLIC 30 30
 FT VARSPPLIC 31 119
 FT VARSPPLIC 148 149
 FT CONFLICT 229 229
 FT CONFLICT 256 258
 FT CONFLICT 270 270
 FT CONFLICT 387 387
 FT CONFLICT 440 440
 FT CONFLICT 508 508
 FT CONFLICT 544 544
 FT CONFLICT 756 756
 FT CONFLICT 765 765
 SQ SEQUENCE 822 AA; 91980 MW; D5A4695FA680926B CRC64;
 Query Match 12.1%; Score 75.5; DB 1; Length 822;
 Best Local Similarity 31.3%; Pred. No. 5.8;
 Matches 26; Conservative 9; Mismatches 31; Indels 17; Gaps 4;
 OY 23 ALGSSVALNCTWVWVSGPRLSPVQWLKGLPLIGCHYSLEHYSWKANLSEVLVSV 82
 DB 268 ALGSSNVEFPCK--VYSDPQ---PHIQWLK-----HIEVNSKIGPDMLPVQILKT 313
 OY 83 LGVNTSTEVYGAFTCSIONISF 105
 DB 314 AGVNTTDKEME---VLHLRNVSF 333
 RESULT 15
 LACH_DROME STANDARD; PRT; 359 AA.
 AC 024372;
 DT 15-JUL-1998 (rel. 36, Created)
 DT 15-JUL-1998 (rel. 36, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Lachesin precursor.
 DE LACHESIN
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94038693; PubMed-8223276;
 RA Karlstrom R.O., Wilder L.P., Bastiani M.J.;
 RT "Lachesin: an immunoglobulin superfamily protein whose expression
 RT correlates with neurogenesis in grasshopper embryos.";
 RT Development 118:509-522(1993).
 CC -1- FUNCTION: MAY PLAY A ROLE IN EARLY NEURONAL DIFFERENTIATION AND

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2002, 17:27:23 : Search time 33 Seconds

(Without alignments)
736.775 Million cell updates/sec

Title: US-09-598-443-2_COPY_1_118

Perfect score: 626

Sequence: 1 MPGVCDRAPEFLSPEDQVL.....SIGNISPSFTLQACPTSH 118

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 671580

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	626	100.0	410	4	Q9H733	Q9H733 homo sapien
2	407.5	65.1	409	11	Q9JLZ8	Q9JLZ8 mus musculu
3	105	16.8	1319	4	Q9HCD3	Q9HCD3 homo sapien
4	96	15.3	534	5	Q9VGS5	Q9VGS5 drosophila
5	94.5	15.1	295	5	Q9VGS29	Q9VGS29 drosophila
6	94	15.0	793	11	Q70246	Q70246 mus musculu
7	93.5	14.9	7962	4	Q10465	Q10465 homo sapien
8	93.5	14.9	34350	4	Q8W242	Q8W242 homo sapien
9	91.5	14.6	1323	13	Q08476	Q08476 gallus gall
10	90	14.4	6620	4	Q96AA2	Q96AA2 homo sapien
11	90	14.4	6831	5	Q23550	Q23550 caenorhabdi
12	90	14.4	7160	5	Q23551	Q23551 caenorhabdi
13	88	14.1	1036	5	Q8SWM3	Q8SWM3 drosophila
14	88	14.1	1336	5	Q9VNI4	Q9VNI4 drosophila
15	88	14.1	2673	4	Q96SC3	Q96SC3 homo sapien
16	88	14.1	5636	4	Q96RW7	Q96RW7 homo sapien

17	87.5	14.0	5198	5	Q76518	Q76518 caenorhabdi
18	84	13.4	298	4	Q96HT1	Q96HT1 homo sapien
19	84	13.4	836	4	Q94856	Q94856 homo sapien
20	84	13.4	1174	11	Q91260	Q91260 rattus norv
21	84	13.4	2053	4	Q8WU7	Q8WU7 homo sapien
22	84	13.4	2113	4	Q8TD84	Q8TD84 homo sapien
23	82.5	13.2	500	5	Q9XZ87	Q9XZ87 drosophila
24	82.5	13.2	4824	5	Q95YM1	Q95YM1 procambarus
25	82.5	13.2	17352	5	Q95YM2	Q95YM2 procambarus
26	82	13.1	582	11	Q8R4B5	Q8R4B5 mus musculu
27	82	13.1	915	11	Q8R4B3	Q8R4B3 mus musculu
28	81.5	13.0	496	5	Q9W260	Q9W260 drosophila
29	81.5	13.0	500	5	Q961W0	Q961W0 drosophila
30	81.5	13.0	650	11	Q63709	Q63709 rattus rat
31	81.5	13.0	1419	13	Q98SM3	Q98SM3 brachydanio
32	81	12.9	1151	11	Q9QVNS	Q9QVNS rattus sp.
33	81	12.9	1217	11	P97685	P97685 rattus norv
34	80.5	12.9	6875	6	Q28733	Q28733 oryctolagus
35	80	12.8	1746	4	Q8WY19	Q8WY19 homo sapien
36	80	12.8	4650	4	Q15598	Q15598 homo sapien
37	80	12.8	6632	5	Q17362	Q17362 caenorhabdi
38	80	12.8	6632	5	Q01761	Q01761 caenorhabdi
39	80	12.8	26926	4	Q10466	Q10466 homo sapien
40	80	12.8	26926	4	Q8W243	Q8W243 homo sapien
41	79.5	12.7	137	13	Q9YHS0	Q9YHS0 ginglymosto
42	79.5	12.7	868	11	Q62838	Q62838 rattus norv
43	79.5	12.7	1256	11	Q9ET59	Q9ET59 mus musculu
44	79.5	12.7	1256	11	Q9J1X1	Q9J1X1 mus musculu
45	79.5	12.7	1256	11	Q92555	Q92555 mus musculu

ALIGNMENTS

RESULT 1

Q9H733 PRELIMINARY: PRT: 410 AA.

ID Q9H733 AC Q9H733.

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE CDNA: FLJ21446 fis, clone COL04458.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-COLON:

RA Kawabata A., Hiki J. T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.,

RT "NEBO human cDNA sequencing project."

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AK025099; BAB15066.1; -

DR InterPro: IPR003600; IG-like.

DR InterPro: IPR000157; TIR-domain.

DR Pfam: PF01582; TIR; 1.

DR PRINTS: PR01559; DUFFYANTIGEN.

DR SMART: SM00410; IG-like; 1.

DR SMART: SM00255; TIR; 1.

SO SEQUENCE 410 AA; 45707 MW; 2A7A663D79567ED6 CRC64;

Query Match 100.0%; Score 626; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 6e-60;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGVCDRAPEFLSPEDQVLRAALGSSVALNTAVVSGPHSLPSVQMLKGLGIGG 60
DB 1 MPGVCDRAPEFLSPEDQVLRAALGSSVALNTAVVSGPHSLPSVQMLKGLGIGG 60
QY 61 HSLTLEYSVKANKLSEVLVSSVLGVNVTSTEVYGAFTCSIGNISPSFTLQACPTSH 118
|||||

Db 61 HYSLEHYSWKANLSEVLVSSVLGVNVTSTVEYGAFTCSIONISFSFTLORAGPTSH 118

RESULT 2.

60JUL28

ID Q9JLZ8 PRELIMINARY: PRT: 409 AA.

AC 01-OCT-2000 (TREMBLrel. 15, Created)

RT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Toll/interleukin-1 receptor 8.

GN SIGIRR OR TIR8.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Pepton-Rol G., Polentarutti N., Muzio M., Mantovani A.;

RT "TIR8: a novel toll/interleukin-1 receptor family member."

EMBL: AF113795, AF26200.1; -

MGI: MGI1344402; Slight.

DR InterPro: IPR003599; Ig.

DR InterPro: IPR003006; TIR_MHC.

DR InterPro: IPR000157; TIR_domain.

DR Pfam: PF00047; Ig: 1.

DR Pfam: PF01582; TIR: 1.

DR SMART: SM00409; IG: 1.

DR SMART: SM00255; TIR: 1.

KW Receptor.

SQ SEQUENCE 409 AA; 45707 MW; DCAAB99A72356A13 CRC64;

Query Match 65.1%; Score 407.5; DB 11; Length 409;

Best Local Similarity 67.8%; Pred. No. 4, 1e-36;

Matches 80; Conservative 8; Mismatches 29; Indels 1; Gaps 1;

QY 1 MPQCDAAPDPLSPSEQVLRPALGSSVALNCTAWVSSPHCSLSYQWLKDLPLGIGG 60

DB 1 MAGVCDAAPNPLSPSEQVLRPALGSSVALNCTAWVSSPHCSLSYQWLKDLGALGNGS 60

QY 61 HYSLEHYSWKANLSEVLVSSVLGVNVTSTVEYGAFTCSIONISFSFTLORAGPTSH 118

DB 61 HFSLEHEDWVANSEI-VSSVLVLMLTNAEDYGTFTCSVMVSSHFTLMRAGPAGH 117

RESULT 3

09HCD3

ID 09HCD3 PRELIMINARY: PRT: 1319 AA.

AC 01-MAR-2001 (TREMBLrel. 16, Created)

RT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE KIAI1639 protein (Fragment).

GN KIAI1639.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEELINE-20450683; Pubmed-10997877;

RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes.

RT XVIII. The complete sequences of 100 new cDNA clones from brain which

RT code for large proteins in vitro."

RL DNA Res. 7:273-281(2000).

EMBL: AB046859; BAB13465.1; -

HSSP: P56276; ITIK

DR InterPro: IPR000719; Euk_Pkinase.

DR InterPro: IPR003961; FN_III.

DR InterPro: IPR001899; Gram_pos_anchor.

DR InterPro: IPR003599; Ig.

RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rehert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klimas I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003556; AAF50456.1; -
DR Flybase: FBgn0035863; CG18650.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 6.
DR SMART: SM00408; IGC2; 4.
KW Immunoglobulin domain.
SQ SEQUENCE 534 AA; 59845 MW; E22071BAE89D2F66 CRC64;
Query Match 15.3%; Score 96; DB 5; Length 534;
Best local similarity 29.6%; Pred. No. 0.055;
Matches 29; Conservative 17; Mismatches 38; Indels 14; Gaps 5;
QY 6 DRAPDLSPSEDQVLRPALGSSVALNCTAWVYSGPHCSLP-SYOMLKDGLPLGIGYSLH 65
DB 156 DASPELLYMSRQTLQP--GPVSLKCYA--TGNP---LPQFWMSLDGFPIDPSRPLVG 208
QY 66 EYSWVANK-LSEVLSSVLCVNTSTEVYGAFTCSION 102
DB 209 QYVTHDVISHVINSV-----KEEDGGEYCTAON 240
RESULT 5
QY9VS29 PRELIMINARY; PRT; 295 AA.
AC Q9VS29; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG8618 protein.
GN CG8618.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandal D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jafali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rehert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klimas I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003560; AAF50601.1; -
DR Flybase: FBgn0035748; CG8618.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 4.
DR SMART: SM00408; IGC2; 2.
KW Immunoglobulin domain.
SQ SEQUENCE 295 AA; 48F45CE162C4FE1 CRC64;
Query Match 15.1%; Score 94.5; DB 5; Length 295;
Best local similarity 28.2%; Pred. No. 0.038;
Matches 29; Conservative 20; Mismatches 39; Indels 15; Gaps 5;
QY 1 MGCVCRRADPFLSPSEDQVLRPALGSSVALNCTAWVYSGPHCSLP-SYOMLKDGLPLGIG 59
DB 91 LPDDIQRLPPLKSLSPFQNTIQLNMGRASLTCS--VVKG---DLPLTIIMRKGRBIDPT 145
QY 60 GHYSIMHSVKNALSEVLSSVLCVNTSTEVYGAFTCSION 102
DB 146 QHMSVKQVD---QYNSILVLENLGSDDT---GMYSCVVRN 179
RESULT 6
QY070246 PRELIMINARY; PRT; 793 AA.
AC 070246; 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative neuronal cell adhesion molecule (PUNC) (Putative neuronal
DE cell adhesion molecule, short form).
GN PUNC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SWISS WEBSTER;
RX MEDLINE=98175891; PubMed=9507132;
RA Salbaum J.M.;
RT "Punc, a novel mouse gene of the immunoglobulin superfamily, is
RT expressed predominantly in the developing nervous system.";
RT Mech. Dev. 71:201-204(1998).
RL [2]
RP SEQUENCE OF 599-715 FROM N.A.
RC STRAIN-FVB;
RX MEDLINE=99122837; PubMed=9922388;
RA Salbaum J.M.;
RT "Genomic structure and chromosomal localization of the mouse gene
RT Punc.";
RL Mamm. Genome 10:107-111(1999).
DR EMBL: AF026465; AAD12133.1; -

```

DR EMBL: AF026466; AAD12124.1; -
DR HSSP: P80362; 1MTL.
DR MGD: MG1:1202390; Punc.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00041; fn3; 2.
DR Pfam: PF00047; fn3; 4.
DR SMART: SM00060; FN3; 2.
DR SMART: SM00408; IG_C2; 4.
DR SMART: SM00408; IG_C2; 4.
DR Immunoglobulin domain.
DR SEQUENCE 793 AA; 84287 MW; 65FD4D1D5E952937 CRC64;

Query Match
Best Local Similarity 15.0%; Score 94; DB 11; Length 793;
Matches 31; Conservative 11; Mismatches 41; Indels 18; Gaps 4;

OY 7 RAPDFLSPEDQVLRPALGSSVALNCTAWVSGPHCSLPVSOMLKDGLPLGIGHYSLHE 66
DB 339 QAPAEVQHPQSTSRPA-GTAMFTQAGQEPPEH-----VTMLKNGVLAGGHVRL-- 390
DB 67 YSWKANLSEVLYSVLVGNVSTVEYGAFTCSIONISFS 107
DB 391 ---KNN-----NSTLSTSGVGPEDALYQCAENIAGSS 421

RESULT 7
ID 010465 PRELIMINARY; PRT; 7962 AA.
AC 01-NOV-1996 (TREMBLrel. 01, Created)
DC 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Titin, skeletal muscle isoform (EC 2.7.1.1-) (Connectin) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_Taxid=9606;
RN [1]_Taxid=9606;
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE.
RX MEDLINE-96026330; PubMed-7569978;
RT Label: S. Kolmer B.;
RT "titin: giant proteins in charge of muscle ultrastructure and
RT elasticity".
RT Science 270:293-296(1995).
CC -1- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.
CC -1- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY
CC ONE TISSUE.
CC -1- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.
CC -1- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE
CC KINASES.
CC -1- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 90
CC IMMUNOGLOBULIN C2-LIKE DOMAINS.
DR EMBL: X90569; CAA62189.1; -
DR HSSP: P56276; 1TLK.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003600; IG_III.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR004168; PPAK_molif.
DR Pfam: PF00047; fn3; 59.
DR Pfam: PF00047; fn3; 59.
DR SMART: SM00060; FN3; 53.
DR SMART: SM00408; IG_C2; 43.
DR SMART: SM00410; IG_Like; 15.
DR Muscle protein: Cytoskeleton: Structural protein; Phosphorylation;
DR Serine/threonine-protein kinase; Alternative splicing; Repeat;
DR Immunoglobulin domain.
FT NON TER 1
DOMAIN 5618 7792 1 GLU/LYS/PRO/VAL-RICH.

```

```

FT NON TER 7962 7962
SO SEQUENCE 7962 AA; 883018 MW; B8524053CBAD58 CRC64;

Query Match
Best Local Similarity 14.9%; Score 93.5; DB 4; Length 7962;
Matches 38; Conservative 7; Mismatches 32; Indels 39; Gaps 7;

OY 10 DFLSPEDQVLR-----PALGSSVALNCTAWVSGPHCSLP-SVQWL 50
DB 1883 DFGSSCDAYLRVDQNDNPSTFKLTKMKDKVLGSSIHMCCK---VSG---SLPISQNF 1936
OY 51 KDGLPLGIGHYSL--HEYSWKANLSEVLYSVLVGNVSTVEYGAFTCSIONIS 104
DB 1937 KDGEIYSASAKRLVCHERS-----VS--LENNNLEDTNNTCKVSNVA 1980

RESULT 8
ID 08WZ42 PRELIMINARY; PRT; 34350 AA.
AC 08WZ42;
DC 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Titin.
GN TN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-20309627; PubMed-10850961;
RX Freiburg A.; Trombittas K.; Hell W.; Cazorla O.; Fougereuse F.;
RA Cantner T.; Kolmerer B.; Wilt C.; Beckmann J.S.; Gregorio C.C.;
RA Granzier H.; Labelt S.;
RT "series of exon-skipping events in the elastic spring region of titin
RT as the structural basis for myofibrillar elastic diversity.".
RT Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE-21573839; PubMed-11717165;
RX Bang M.L.; Centner T.; Fornoff F.; Geach A.J.; Gotthardt M.;
RA McNabb M.; Wilt C.C.; Labelt D.; Gregorio C.C.; Granzier H.;
RA Labelt S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT 2-1 line to I-band linking system.".
RT Circ. Res. 89:1065-1072(2001).
DR EMBL: AJ277892; CAD12456.1; -
DR InterPro: IPR000282; CytoK_receptor_2.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR000577; FGGL_Kin.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR001092; HLM_basic.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR InterPro: IPR002016; Peroxidase.
DR InterPro: IPR004168; PPAK_molif.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00041; fn3; 132.
DR Pfam: PF00047; fn3; 146.
DR Pfam: PF00069; kinase; 1.
DR Pfam: PF02818; PPAK; 53.
DR Pfam: PF00001; Euk_pkinase; 1.
DR SMART: SM00060; FN3; 133.
DR SMART: SM00409; IG; 167.
DR SMART: SM00408; IG_C2; 148.
DR SMART: SM00406; IGV; 23.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TYRKC; 1.

```


RX MEDLINE=3387664; PubMed=83971735.
 RA Banian G.M., L'Hennault S.W., Morris M.E.;
 RT "Additional sequence complexity in the muscle gene, unc-22, and its
 RT encoded protein, twitchin, of *Caenorhabditis elegans*.";
 RL Genetics 134:1097-1104(1993)
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS: ISOFORM ZK617.1A (Q23550) AND
 CC ISOFORM ZK617.1B (Q23551); MAY BE PRODUCED BY ALTERNATIVE
 CC SPLICING
 CC
 GC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: Z73893; CAA8804.1; JOINED.
 DR EMBL: Z73893; CAA8804.1; JOINED.
 DR EMBL: Z73893; CAA8808.1; JOINED.
 DR EMBL: X15420; CAA33463.1; -.
 DR HSSP: Q63450; IA06.
 DR WormPep: ZK617.1A; CE06634.
 DR Interpro: IPR00719; Euk_Pkinase.
 DR Interpro: IPR003962; FNIII_repeat.
 DR Interpro: IPR003961; FN_III.
 DR Interpro: IPR003598; Ig_c2.
 DR Interpro: IPR003600; Ig_Like.
 DR Interpro: IPR003006; Ig_MHC.
 DR Interpro: IPR002290; Ser_thr_kinase.
 DR Interpro: IPR001412; tRNA-synt_I.
 DR Pfam: PF00041; fn3; 31.
 DR Pfam: PF00047; Ig; 22.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00014; ENTPEPITL.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00060; FN3; 27.
 DR SMART: SM00408; IGC2; 6.
 DR SMART: SM00410; IG_Like; 17.
 DR SMART: SM00220; STC; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 DR APP-Binding: Alternative splicing, Hypothetical protein;
 KM Immunoglobulin domain; Repeat; Serine/threonine-protein kinase;
 KM Transferrase
 SQ SEQUENCE 6831 AA: 752596 MW: 6CADE032963A52E1 CRC64;
 Query Match 14.4% Score 90: DB 5; Length 6831;
 Best Local Similarity 24.0%; Pred. NO. 5.6;
 Matches 23; Conservative 21; Mismatches 40; Indels 12; Gaps 3;
 QY 7 RAEPLSPSEDOYLRLPALGSSVALNCTAMVVSQPHCSLPVSOMLKDGLPIGIGHXSLHE 66
 Db 101 RGSSFV--GGPRRIIPDGAGALIVMECKY-----KSASPPVAKKMKMGDGVPLSMGGLYHA-- 151
 Db 67 YSWKANLESVLVSSVLGVNVSTFVGATFCISION 102
 Db 152 ---IFSDLDGDTYLTCLLETIRGPSSSDAQGYRCNTRN 184
 RESULT 12
 ID Q23551 PRELIMINARY; PRT; 7160 AA.
 AC Q23551;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE UNC-22 protein.
 DE UNC-22 OR ZK617.1
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdittida; Rhabditoidea;
 OC Rhabdillidae; Pezodermatidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harris B., White S.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 OG -1- ALTERNATIVE PRODUCTS: TWO FORMS: ISOFORM ZK617.1A (Q23550) AND

CC ISOFORM SK617.1B (SHOWN HERE); MAY BE PRODUCED BY ALTERNATIVE
CC SPLICING.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASRS.
DR EMBL; Z73899; CAA98082.1; -. JOINED.
DR EMBL; Z73897; CAA98082.1; -. JOINED.
DR EMBL; Z73899; CAA98085.1; -. JOINED.
DR HSSP; O63450.1A06. JOINED.
DR WORMPEP; ZK617.1B; CE06635.
DR INTERPRO; IPR000719; Euk_Pkinase.
DR INTERPRO; IPR003962; Fm11_repeat.
DR INTERPRO; IPR003961; Fm_11t.
DR INTERPRO; IPR003958; Ig_c2.
DR INTERPRO; IPR003600; Ig_Like.
DR INTERPRO; IPR002290; Ser_thr_kinase.
DR INTERPRO; IPR001412; tRNA-synt_1.
DR Pfam; PF00047; fn3; 31.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00014; FNTPR11.
DR PRODOM; PD000001; Euk_Pkinase; 1.
DR SMART; SM00060; FN3; 27.
DR SMART; SM00408; IGc2; 6.
DR SMART; SM00410; IG_Like; 18.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Alternative splicing; Hypothetical protein;
KW Immunoglobulin domain; Repeat; Serine/threonine-protein kinase;
KW Transferrase

SQ SEQUENCE 7160 AA: 789230 MW: 387D0M4F6Z7CD027 CRC64;

Query Match 14.4%; Score 90; DB 5; Length 7160;
Best Local 23; Conservativity 21.08; Pred. NO. 6;
Matches 23; Conservative 21; Mismatches 40; Indels 12; Gaps 3;

OY 7 RAPDFLSSEDDVLRPALGSSVALNCTIWMVVSGPCSLPSYQWLKDGLPIGCHYSHE 66
Db REPSTV--GKRPIIPKDGALIVMCKV-----KSASTPVAKMKMDGVLSMGLYHA-- 151
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| :
OY 67 YSWKANLEVALSVSLGCVNTSTEVEYGATFCOSION 102
Db 152 --IFSDLGDDTYTLQLLEIRGPSSSDAQYRCNIKN 184
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| :
| :

RESULT 13

OBSMW3 PRELIMINARY; PTR; 1036 AA.

AC OBSMW3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
ID LD28757P.
GN CG1084.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agabany A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA Miranda R., Gonzalez M., Guarin H., Kromliller B., Li P., Liao G.,
RA Andrade A., Mangalaj C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Cataliker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AY095040; AAM11368.1; .

SO SEQUENCE 1036 AA; 117742 MW; A370F297D39D719D CRC64;
Query Match 14.1%; Score 88; DB 5; Length 1036;
Best Local Similarity 27.7%; Pred. No. 0.92;
Matches 28; Conservative 9; Mismatches 42; Indels 22; Gaps 4;
OY 4 VCDRAPDPLS-PSEDOVLRPALGSSVALNCTAWVSGPHCS-LPSVQWLKDLPLIGIGH 61
DB 397 VLSMKPSFKKHPLSESVY-----AVYNGNTTIVCDPEAPRPFKQMKKDGQVIGSGH 449
OY 62 YSLHEYSWKANLSEVLSVGVNTSTEVGATCSTQN 102
DB 450 -----RRLPSGTLTISPTSRDDEGIYTCIASN 477
RESULT 14
O9VN14 PRELIMINARY; PRT; 1336 AA.
O9VN14: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG1084 protein.
GN CG1084.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.B., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster."
DR EMBL; AF003606; AAF52137.1; -
DR HSSP; P20241; 1CFB.

DR FLYBase; FBgn0037240; CG1084.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR003598; Ig.C2.
DR InterPro; IPR003600; Ig_1like.
DR InterPro; IPR003006; Ig_1MHC.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00041; fn3; 4.
DR Pfam; PF00047; Ig; 5.
DR Pfam; PF00059; Lectin_C; 1.
DR SMART; SM00034; CLEC; 1.
DR SMART; SM00060; FN3; 4.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00410; IG_1like; 2.
DR PROSITE; PS0041; C_Type_Lectin_2; 1.
KW Immunoglobulin domain.
SQ SEQUENCE 1336 AA; 151994 MW; 1636094A64484262 CRC64;
Query Match 14.1%; Score 88; DB 5; Length 1336;
Best Local Similarity 27.7%; Pred. No. 1.3;
Matches 28; Conservative 9; Mismatches 42; Indels 22; Gaps 4;
OY 4 VCDRAPDPLS-PSEDOVLRPALGSSVALNCTAWVSGPHCS-LPSVQWLKDLPLIGIGH 61
DB 697 VLSMKPSFKKHPLSESVY-----AVYNGNTTIVCDPEAPRPFKQMKKDGQVIGSGH 749
OY 62 YSLHEYSWKANLSEVLSVGVNTSTEVGATCSTQN 102
DB 750 -----RRLPSGTLTISPTSRDDEGIYTCIASN 777
RESULT 15
O96SC3 PRELIMINARY; PRT; 2673 AA.
O96SC3: 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fibulin-6 (Fragment).
GN FIBL-6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Carnivora; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Kostka G., Timpi R.;
RT "Partial sequence of fibulin-6 with a c-terminal region related to domain II and III of the fibulin family."
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ306906; CAC37630.1; -
DR InterPro; IPR00152; Asx_hydroxyl.
DR InterPro; IPR000875; Cecropin.
DR InterPro; IPR000561; EGF_1like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR003006; Ig_1MHC.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF00047; Ig; 17.
DR Pfam; PF00090; TSP_1; 6.
DR PROSITE; PS0010; ASX_HYDROXYL; UNKNOWN_5.
DR PROSITE; PS00268; CECROPIN; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_8.
DR PROSITE; PS00092; TSP1; 6.
FT NON_TER
SQ SEQUENCE 2673 AA; 291017 MW; BEAEC30B8340E272 CRC64;
Query Match 14.1%; Score 88; DB 4; Length 2673;
Best Local Similarity 28.6%; Pred. No. 2.9;
Matches 26; Conservative 13; Mismatches 32; Indels 20; Gaps 4;
OY 12 LSPSEDVLRPALGSSVALNCTAWVSGPHCS-LPSVQWLKDLPLIGIGHYSLHEYSWK 71

